

RESEARCH OUTPUTS / RÉSULTATS DE RECHERCHE

Géométrie Symplectique, BAC3 mathématique

Carletti, Timoteo; Libert, Anne-Sophie

Publication date:
2009

Document Version
Early version, also known as pre-print

[Link to publication](#)

Citation for published version (HARVARD):
Carletti, T & Libert, A-S 2009, *Géométrie Symplectique, BAC3 mathématique..*

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Table des matières

<i>Avant-propos.</i>	<i>iii</i>
Chapitre I. <i>Rappels de mécanique lagrangienne.</i>	<i>1</i>
1. La mécanique Lagrangienne	1
2. Méthodes variationnelles en mécanique Lagrangienne	4
2.1. Les équations d'Euler	6
2.2. Le plus court chemin entre deux points	8
2.3. Principe variationnel d'Hamilton : formulation Lagrangienne	9
Chapitre II. <i>Introduction à la mécanique hamiltonienne.</i>	<i>11</i>
1. La transformation de Legendre	11
2. La fonction d'Hamilton	14
3. Les équations d'Hamilton	15
4. Théorème de Liouville	16
5. Théorème de Poincaré : Récurrence	19
6. Méthodes variationnelles en mécanique Hamiltonienne	19
6.1. Principe de l'action stationnaire ou de Maupertuis	21
Chapitre III. <i>Formalisme canonique.</i>	<i>23</i>
1. Introduction	23
2. Structure symplectique de l'espace des phases.	24
3. Champs Hamiltoniens	28
4. Transformations canoniques	30
5. Fonctions, champs de vecteurs et formes différentielles sur variétés	37
5.1. Fonctions	37
5.2. Champs de vecteurs	37
5.3. 1-formes différentielles	39
5.4. 2-formes différentielles	40
5.5. Quelques définitions	41
6. Condition de Lie	42
7. Fonctions génératrices (de transformations canoniques)	46
8. Parenthèses de Poisson	49
8.1. Parenthèses de Poisson et intégrales premières	50
8.2. Parenthèses de Poisson et transformations canoniques	51
9. Invariants intégraux	52
Chapitre IV. <i>Transformations proches de l'identité : Série de Lie.</i>	<i>55</i>
1. Introduction	55
Chapitre V. <i>Théorie d'Hamilton–Jacobi et variables Action–Angle 1–dim.</i>	<i>59</i>
1. Introduction et exemples	59

p. ii	Chapitre
2. Séparation des variables	62
3. Variables Angle–Action en dimension 1	65
Chapitre VI. <i>Intégration par quadratures : Liouville, Arnol’d et Action–Angle.</i>	73
1. Introduction	73
2. Intégration par quadrature : le Théorème de Liouville	74
3. Le Théorème d’Arnol’d	78
4. Systèmes hamiltoniens avec plusieurs degrés de liberté : variables action–angle	79
5. Mouvements quasi–périodiques	82
5.1. $l = 1$: dynamique sur \mathbb{T}	82
5.2. $l = 2$: dynamique sur \mathbb{T}^2	82
5.3. $l \geq 3$: dynamique sur \mathbb{T}^l	83
Chapitre VII. <i>Bibliographie.</i>	91

Avant-propos.

Ces notes ont été rédigées pour la préparation du cours de *Géométrie Symplectique* de troisième année BAC3, que j’ai donné au Département de Mathématique Université Notre Dame de la Paix – Namur, Belgique, dans la période février–mai 2007. Le matériel a été ensuite amélioré, corrigé et complété grâce à l’aide des assistants, que je tiens à remercier très chaleureusement, en particulier Anne-Sophie Libert pour l’esprit critique et minutieux de ses relectures.

Les étudiants pourront y trouver réunie, travaillée et adaptée à un cours de 22.5 heures, la matière prise de plusieurs “sources”, citées en bibliographie, auxquelles je renvoie pour tout renseignement complémentaire et approfondissement.

La géométrie symplectique est un domaine qui se trouve à l’intersection de plusieurs branches des mathématiques, notamment la géométrie différentielle, l’analyse, les systèmes dynamiques, l’algèbre. J’ai cependant décidé, soit pour des raisons de recherche personnelle, soit pour bien insérer ce cours dans le contexte des cours proposés au Département de mathématique, de lui donner plutôt une vue “systèmes dynamiques”, en mettant en évidence certains aspects de la théorie des systèmes Hamiltoniens, et passant en revue rapidement d’autres parties plus liées à la géométrie et l’algèbre. Par nature donc, ces notes seront “incomplètes” et j’invite les étudiants intéressés à les compléter avec la bibliographie proposée.

A cause du “public visé”, ces notes ne demandent pas beaucoup de prérequis ; avoir suivi un cours de mécanique du point et mécanique Lagrangienne sera normalement suffisant pour la partie systèmes Hamiltoniens. Des connaissances de géométrie différentielle (au moins les surfaces) et quelques notions d’algèbre (formes algébrique et formes différentielles) seront suffisantes pour la partie dédiée à la géométrie symplectique.

Nous terminons cet avant-propos avec une remarque concernant la numération des équations, sections, sous-sections etc. Les équations sont numérotées consécutivement à l’intérieur de chaque section et leur numération contient toujours le nombre du Chapitre, c’est-à-dire équation (I.2.3) signifie la troisième équation de la seconde section du premier Chapitre. Définitions, Exemples, Remarques, Théorèmes, Corollaires, Lemmes et tout autre objet seront numérotés consécutivement à l’intérieur de chaque section, leur numération ne contient pas le numéro du Chapitre, mais seulement celui de la section. On aura donc Définition 3.1, le premier objet de la section 3, et qui sera suivie par Remarque 3.2, et cetera.

Les références aux objets (autres que équations) contiennent le numéro du Chapitre si elle font référence à un objet contenu dans un Chapitre différent du Chapitre courant. Donc objet I.2.3 signifie le troisième objet de la seconde section du premier Chapitre, si on n’est pas en train de lire le premier Chapitre, autrement le même objet sera indiqué par 2.3 dans le Chapitre I.

TIMOTEO CARLETTI
DÉPARTEMENT DE MATHÉMATIQUE FUNDP
REMPART DE LA VIERGE, 8
B-5000 NAMUR BELGIQUE
TIMOTEO.CARLETTI@FUNDP.AC.BE

Namur, Janvier 2008

premières corrections Septembre 2007

deuxièmes corrections Janvier 2008

CHAPITRE I

Rappels de mécanique lagrangienne.

1. La mécanique Lagrangienne

La mécanique Lagrangienne permet de traiter, et parfois de résoudre, des problèmes mécaniques où plusieurs (mais toujours en nombre fini) points interagissent entre eux avec des forces (contraintes, forces de liaison). Cela est essentiellement dû au fait que l'équation de Newton

$$\vec{F} = m\vec{a},$$

demande de résoudre $3n$ équations différentielles du deuxième ordre ; par contre l'introduction et l'usage des *coordonnées généralisées* permet de réduire le nombre d'équations à résoudre.

Par exemple dans le cas de *liaisons holonomes*, c'est-à-dire indépendantes du temps, et *sans friction*, un système de points matériels (chaque point P_i ayant une masse $m_i > 0$) sur lesquels agissent des forces extérieures provenant d'une énergie potentielle V , $\vec{F} = -\vec{\nabla}V$, peut être décrit par une *fonction de Lagrange*, ou *Lagrangienne* :

$$(I.1.1) \quad \mathcal{L}(\vec{q}, \dot{\vec{q}}) = T - V,$$

où $\vec{q} = (q_1, \dots, q_m)$ est le vecteur des *coordonnées généralisées* ou *lagrangiennes*, $\dot{\vec{q}}$ sa dérivée par rapport au temps t . T représente l'*énergie cinétique* et comme déjà dit V l'*énergie potentielle*. L'espace à $2m$ dimensions des couples de vecteurs $(\vec{q}, \dot{\vec{q}})$ est appelé *espace des configurations*, chaque "point" dans cet espace détermine de façon univoque la position et la vitesse instantanée du système.

L'évolution, c'est-à-dire la dynamique, du système des points est obtenue en résolvant les équations de *Euler-Lagrange* :

$$(I.1.2) \quad \forall j = 1, \dots, m \quad \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \frac{\partial \mathcal{L}}{\partial q_j} = 0.$$

REMARQUE 1.1. Si l'énergie cinétique, comme c'est souvent le cas, est une fonction quadratique des vitesses généralisées, $\dot{\vec{q}}$, alors les équations de Euler-Lagrange sont des équations différentielles du second ordre, mais en nombre réduit par rapport au nombre des points matériels composant le système.

En fait :

$$T = \frac{1}{2} \sum_{i,j=1}^m a_{ij} \dot{q}_i \dot{q}_j + \sum_{i=1}^m b_i \dot{q}_i + c,$$

où la matrice $A = (a_{ij})$ est définie positive et symétrique. Donc de (I.1.2) on obtient pour tout $i = 1, \dots, m$:

$$\frac{d}{dt} \left[\sum_{j=1}^m a_{ij} \dot{q}_j + b_i \right] - \frac{\partial \mathcal{L}}{\partial q_i} = 0,$$

et finalement :

$$\sum_{j=1}^m a_{ij} \ddot{q}_j = F(\vec{q}, \dot{\vec{q}}).$$

Terminons cette introduction avec deux exemples de mécanique Lagrangienne.

EXEMPLE 1.2 (cercle et ressort). *Considérons deux points de masses respectivement m_1 et m_2 contraints de rester sur un plan vertical. Le point P_1 est attaché à un rail circulaire lui aussi disposé sur le même plan vertical et P_2 sur une droite horizontale toujours dans le même plan. Chaque point est soumis à la force de gravité verticale et les points sont reliés par un ressort (voir Fig. 1).*

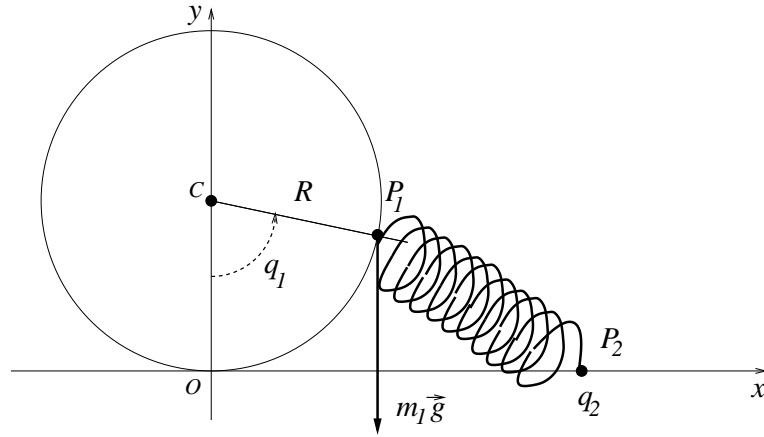


FIG. 1

En coordonnées cartésiennes les points peuvent être repérés par :

$$P_1 = (x_1, y_1) \quad \text{et} \quad P_2 = (x_2, y_2),$$

les contraintes sont déterminées par :

$$x_1^2 + y_1^2 - 2Ry_1 = 0 \quad \text{et} \quad y_2 = 0.$$

Nous pouvons introduire des coordonnées généralisées q_1 et q_2 comme montré dans la Fig. 1 :

$$x_1 = R \sin q_1, \quad y_1 = R - R \cos q_1 \quad \text{et} \quad x_2 = q_2,$$

sous forme vectorielle $\vec{q} = (q_1, q_2)$.

L'énergie cinétique est :

$$T(\vec{q}, \dot{\vec{q}}) = \frac{m_1}{2} R^2 \dot{q}_1^2 + \frac{m_2}{2} \dot{q}_2^2,$$

et l'énergie potentielle est donnée par la somme de deux termes : un terme dû à la gravitation et un terme dû au ressort (énergie élastique) :

$$V_{grav}(\vec{q}) = m_1 g R (1 - \cos q_1) \text{ et } V_{el}(\vec{q}) = \frac{k}{2} (q_2^2 + 2R^2 - 2Rq_2 \sin q_1 - 2R^2 \cos q_1).$$

Puisque le point P_2 reste sur l'axe horizontal son énergie gravitationnelle est une constante qui peut être retirée, autrement dit la force de pesanteur ne fait pas de travail car la contrainte est supposée réagir de façon idéale.

Le Lagrangien du système est donc :

$$\mathcal{L}(\vec{q}, \dot{\vec{q}}) = \frac{m_1}{2} R^2 \dot{q}_1^2 + \frac{m_2}{2} \dot{q}_2^2 - m_1 g R (1 - \cos q_1) - \frac{k}{2} (q_2^2 + 2R^2 - 2Rq_2 \sin q_1 - 2R^2 \cos q_1),$$

et les équations d'Euler-Lagrange :

$$\ddot{q}_1 + \frac{g}{R} \sin q_1 - \frac{k}{m_1} \left(\frac{q_2}{R} \cos q_1 - \sin q_1 \right) = 0 \text{ et } \ddot{q}_2 + \frac{k}{m_2} (q_2 - R \sin q_1) = 0.$$

EXEMPLE 1.3 (Force centrale, problème de Kepler). Considérons un point de masse m soumis à une force centrale $\vec{F}(\vec{r}) = f(r)\vec{r}/r$, où $\vec{r} = (x, y, z)$ est le vecteur position dans \mathbb{R}^3 , $r = \sqrt{x^2 + y^2 + z^2}$ est son module et $f(r)$ est une fonction qui détermine le type de force que nous voulons considérer. Supposons que la fonction $f(r)$ soit régulière, alors on peut déterminer une énergie potentielle :

$$V(r) = - \int_{r_0}^r f(\rho) d\rho.$$

Le moment angulaire par rapport à l'origine est :

$$\vec{L} = \vec{r} \times m\dot{\vec{r}},$$

il est donc clair que le moment angulaire est préservé (il est une intégrale première du mouvement), car :

$$\frac{d}{dt} \vec{L} = \frac{d}{dt} \vec{r} \times m\dot{\vec{r}} + \vec{r} \times m\ddot{\vec{r}} = \vec{0} + \vec{r} \times f(r) \frac{\vec{r}}{r} = \vec{0}.$$

Ce qui implique, si $\vec{L} \neq 0$, que le mouvement sera restreint à un plan, le plan dont la normale est \vec{L} . Nous allons dans la suite supposer ce cas. Si par contre $\vec{L} = 0$ alors le mouvement sera sur une droite (car \vec{r} est parallèle à $\dot{\vec{r}}$).

Puisque le mouvement se déroule entièrement sur un plan, quitte à faire une rotation nous pouvons prendre les coordonnées rectangulaires (x, y) sur ce plan pour repérer le point P et ensuite des coordonnées polaires : $x = r \cos \phi$ et $y = r \sin \phi$ (voir Fig. 2).

L'énergie cinétique est :

$$T = \frac{m}{2} (\dot{r}^2 + r^2 \dot{\phi}^2),$$

et l'énergie potentielle :

$$V = V(r),$$

par exemple si on considère la force d'attraction Newtonienne, $\vec{F}(\vec{r}) = -\frac{k}{r^2} \frac{\vec{r}}{r}$, alors $V(r) = -\frac{k}{r}$.

Le Lagrangien du système résultant :

$$\mathcal{L}(\vec{r}, \dot{\vec{r}}) = \frac{m}{2} (\dot{r}^2 + r^2 \dot{\phi}^2) - V(r),$$

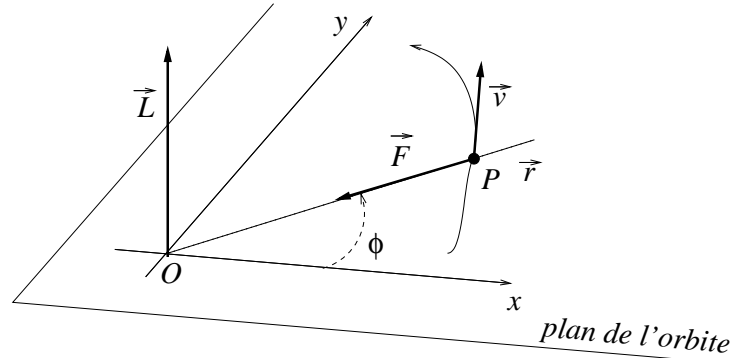


FIG. 2

et les équations de Lagrange :

$$m\ddot{r} - rm\dot{\phi}^2 + \frac{dV(r)}{dr} = 0 \quad \text{et} \quad m \frac{d}{dt}(r^2\dot{\phi}) = 0.$$

Dans la section suivante nous montrerons que les équations d'Euler-Lagrange peuvent être obtenues dans un formalisme beaucoup plus général que celui de la mécanique (déplacements virtuels).

2. Méthodes variationnelles en mécanique Lagrangienne

Dans cette section nous allons introduire les *méthodes variationnelles* (dans le formalisme Lagrangien) et nous montrerons leur lien étroit avec les équations d'Euler-Lagrange. Les premières études sur le *calcul des variations* ont été faites par Euler au XVII^e siècle, c'est à lui qu'on doit un des Théorèmes fondamentaux de la théorie.

Les méthodes variationnelles sont motivées par le fait suivant : si avec plusieurs possibilités un phénomène (physique) se réalise, c'est parce qu'il rend *stationnaire* une "certaine fonction". En mécanique cela peut être relu comme suit : le *mouvement naturel*, c'est-à-dire celui qui découle des équations d'Euler-Lagrange, est celui qui rend stationnaire une certaine fonctionnelle (une fonction dont les arguments sont des fonctions) dans la classe de tous les mouvements *a priori* possibles avec certaines contraintes. En d'autres termes nous verrons que les solutions des équations d'Euler-Lagrange ont cette propriété de minimiser une fonctionnelle ; mais on ira plus loin car on pourra montrer que tout élément qui minimise cette fonctionnelle sera une solution des équations d'Euler-Lagrange, et donc un mouvement naturel. C'est pour cela que nous pouvons parler de *principes variationnels*, nous pourrions, contrairement à ce que nous avons fait, supposer (*principe de la mécanique classique*) que le mouvement naturel est celui qui minimise la fonctionnelle, et donc les équations d'Euler-Lagrange en découleraient naturellement. Il s'avère que cette façon de procéder est très utile dans beaucoup de situations, notamment en relativité, en théorie de champs, en mécanique statistique.

Commençons par un simple exemple.

EXEMPLE 2.1 (Mouvement rectiligne uniforme). *Considérons un point de masse m sur lequel n'agit aucune force et qui se déplace dans \mathbb{R}^3 .*

Nous savons (équations de Newton ou Euler–Lagrange) que son mouvement sera une ligne droite et que sa vitesse sera constante. Si au temps t_0 le point se trouve en \vec{x}_0 alors pour tout t :

$$\vec{x}_{nat}(t) = \vec{x}(t_0) + \vec{v}(t_0)(t - t_0) .$$

Fixons un intervalle de temps $[t_0, t_1]$ pour un quelconque $t_1 > t_0$; nous allons montrer que le mouvement naturel minimise la fonctionnelle énergie cinétique :

$$\mathcal{F}(\vec{y}, \dot{\vec{y}}) := \int_{t_0}^{t_1} \frac{m}{2} \dot{\vec{y}}^2(s) ds ,$$

dans la classe de toutes les perturbations avec les mêmes conditions initiales. Autrement dit pour aller d'un point P_1 à un deuxième point P_2 un corps libre choisira de parcourir parmi tous les chemins possibles celui qui minimise l'énergie cinétique.

Considérons donc la famille suivante de mouvements variés, H , (voir Fig. 3) :

$$(I.2.1) \quad \vec{x}_\eta(t) = \vec{x}_{nat}(t) + \vec{\eta}(t) \quad \forall t_0 \leq t \leq t_1 ,$$

où $\vec{\eta}$ est une perturbation de classe \mathcal{C}^2 telle que :

$$\vec{\eta}(t_0) = \vec{\eta}(t_1) = 0 .$$

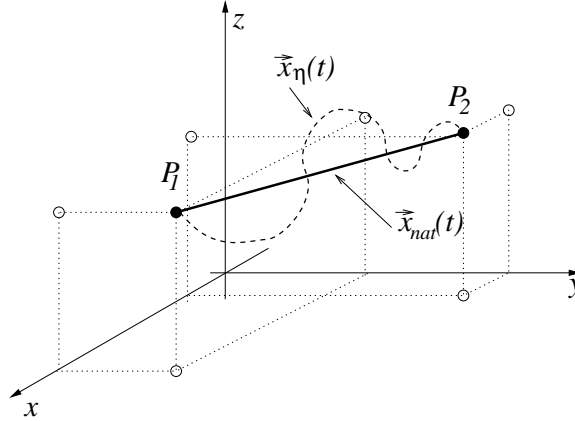


FIG. 3

Calculons la variation de la fonctionnelle \mathcal{F} , c'est-à-dire la différence de la fonctionnelle calculée sur le mouvement naturel et sur le mouvement varié (nous pouvons penser à une dérivée mais faite sur un espace de dimension infinie) :

$$\delta\mathcal{F}(\vec{x}_{nat}) := \mathcal{F}(\vec{x}_\eta, \dot{\vec{x}}_\eta) - \mathcal{F}(\vec{x}_{nat}, \dot{\vec{x}}_{nat}) = \int_{t_0}^{t_1} \frac{m}{2} \left(\dot{\vec{x}}_\eta^2(s) - \dot{\vec{x}}_{nat}^2(s) \right) ds .$$

Si nous introduisons la définition du mouvement varié nous avons :

$$\begin{aligned} \delta\mathcal{F}(\vec{x}_{nat}) &= \int_{t_0}^{t_1} \frac{m}{2} \left(\dot{\vec{x}}_{nat}^2(s) + 2\dot{\vec{x}}_{nat}(s) \cdot \dot{\vec{\eta}}(s) + \dot{\vec{\eta}}^2(s) - \dot{\vec{x}}_{nat}^2(s) \right) ds \\ &= \int_{t_0}^{t_1} \frac{m}{2} \left(2\dot{\vec{x}}_{nat}(s) \cdot \dot{\vec{\eta}}(s) + \dot{\vec{\eta}}^2(s) \right) ds , \end{aligned}$$

Le terme en $\dot{\vec{\eta}}$ peut être éliminé par une intégration par partie et en utilisant (I.2.1) (c'est-à-dire $\ddot{\vec{x}}_{nat} \equiv 0$) :

$$\int_{t_0}^{t_1} m\dot{\vec{x}}_{nat}(s) \cdot \dot{\vec{\eta}}(s) ds = m\dot{\vec{x}}_{nat}(s) \cdot \vec{\eta}(s) \Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} m\ddot{\vec{x}}_{nat}(s) \cdot \vec{\eta}(s) ds = 0,$$

car $\ddot{\vec{x}}_{nat} \equiv 0$ et $\vec{\eta}(t_0) = \vec{\eta}(t_1) = 0$. Donc

$$\delta\mathcal{F}(\vec{x}_{nat}) = \int_{t_0}^{t_1} \frac{m}{2} \dot{\vec{\eta}}^2(s) ds,$$

et nous pouvons donc conclure que (rappelez-vous que la perturbation est de classe \mathcal{C}^2) :

$$\delta\mathcal{F}(\vec{x}_{nat}) \geq 0 \quad \forall \vec{\eta} \in H,$$

et

$$\delta\mathcal{F}(\vec{x}_{nat}) = 0 \Leftrightarrow \vec{\eta} \equiv 0,$$

c'est-à-dire le mouvement naturel réalise le minimum.

2.1. Les équations d'Euler. Considérons maintenant la situation générale suivante. Soit $F : \mathbb{R}^{2m+1} \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^2 et soit :

$$(I.2.2) \quad \mathcal{Q} = \{\vec{q} : \mathbb{R} \rightarrow \mathbb{R}^m : \vec{q} \in \mathcal{C}^2[t_0, t_1], \vec{q}(t_0) = \vec{q}_0 \text{ et } \vec{q}(t_1) = \vec{q}_1\},$$

pour deux “temps” t_0 et t_1 et deux points \vec{q}_0 et \vec{q}_1 dans \mathbb{R}^m fixés. \mathcal{Q} est l'ensemble de toutes les courbes \mathcal{C}^2 qui sont en \vec{q}_0 au temps t_0 et en \vec{q}_1 au temps t_1 .

Introduisons la fonctionnelle :

$$(I.2.3) \quad \mathcal{F}(\vec{q}, \dot{\vec{q}}) := \int_{t_0}^{t_1} F(\vec{q}(s), \dot{\vec{q}}(s), s) ds,$$

et cherchons les conditions qui rendent la fonctionnelle *stationnaire*. Pour faire cela nous avons besoin d'introduire l'équivalent des mouvements variés ; soit donc \vec{q}_* un élément de \mathcal{Q} , définissons l'ensemble des perturbations :

$$(I.2.4) \quad H := \{\vec{\eta} : \mathbb{R} \rightarrow \mathbb{R}^m : \vec{\eta} \in \mathcal{C}^2[t_0, t_1], \vec{\eta}(t_0) = \vec{\eta}(t_1) = 0\},$$

et considérons pour $\vec{\eta} \in H$ fixé, le sous-ensemble $\mathcal{Q}_\eta \subset \mathcal{Q}$ des courbes telles que chaque composante peut s'écrire sous la forme :

$$(I.2.5) \quad q_k(t) = q_{*,k}(t) + \alpha_k \eta_k(t) \quad \forall k = 1, \dots, m \quad \forall t \in [t_0, t_1],$$

où le vecteur $\vec{\alpha} = (\alpha_1, \dots, \alpha_m)$ appartient à \mathbb{R}^m . La restriction de \mathcal{F} aux courbes dans \mathcal{Q}_η est une fonction des variables $(\alpha_k)_{1 \leq k \leq m}$. Nous dirons que la fonctionnelle \mathcal{F} est *stationnaire* dans \mathcal{Q} en \vec{q}_* si la restriction $\mathcal{F}_\eta(\vec{\alpha}) = \mathcal{F}(\vec{q}_* + \vec{\alpha}\vec{\eta}, \dot{\vec{q}}_* + \vec{\alpha}\dot{\vec{\eta}})$ est stationnaire (comme fonction de m variables) pour $\vec{\alpha} = 0$ et pour tout $\vec{\eta} \in H$.

Le résultat suivant, dû à Euler, donne des conditions nécessaire et suffisantes pour avoir la condition de fonctionnelle stationnaire.

THÉORÈME 2.2. Une condition nécessaire et suffisante pour que la fonctionnelle \mathcal{F} soit stationnaire dans \mathcal{Q} en \vec{q}_* est que les composantes $(q_{*,k}(t))_{1 \leq k \leq m}$ soient solutions des équations d'Euler :

$$(I.2.6) \quad \frac{d}{dt} \frac{\partial F}{\partial \dot{q}_k} - \frac{\partial F}{\partial q_k} = 0.$$

DÉMONSTRATION. Introduisons la forme (I.2.5) dans la définition de la fonctionnelle (pour un $\vec{\eta} \in H$ quelconque) et dérivons par rapport aux variables $(\alpha_k)_{1 \leq k \leq m}$ (les conditions de régularité assurent qu'on puisse dériver à l'intérieur de l'intégrale) :

$$\begin{aligned} \frac{\partial}{\partial \alpha_k} \mathcal{F}_\eta(\vec{\alpha}) &= \int_{t_0}^{t_1} \sum_{l=1}^m \left[\frac{\partial F}{\partial q_l} \frac{\partial(q_{*,l}(s) + \alpha_l \eta_l(s))}{\partial \alpha_k} + \frac{\partial F}{\partial \dot{q}_l} \frac{\partial(\dot{q}_{*,l}(s) + \alpha_l \dot{\eta}_l(s))}{\partial \alpha_k} \right] ds \\ &= \int_{t_0}^{t_1} \left[\frac{\partial F}{\partial q_k} \eta_k(s) + \frac{\partial F}{\partial \dot{q}_k} \dot{\eta}_k(s) \right]_{\vec{q}_* + \vec{\alpha} \vec{\eta}} ds, \end{aligned}$$

Si maintenant nous imposons $\vec{\alpha} = 0$ alors nous obtenons :

$$(I.2.7) \quad \frac{\partial}{\partial \alpha_k} \mathcal{F}_\eta(0) = \int_{t_0}^{t_1} \left[\frac{\partial F}{\partial q_k} \eta_k(s) + \frac{\partial F}{\partial \dot{q}_k} \dot{\eta}_k(s) \right]_{\vec{q}=\vec{q}_*} ds.$$

Transformons maintenant le deuxième terme avec une intégration par partie et rappelons-nous la définition des perturbations H (cfr. (I.2.4)) :

$$\begin{aligned} \frac{\partial}{\partial \alpha_k} \mathcal{F}_\eta(0) &= \int_{t_0}^{t_1} \left[\frac{\partial F}{\partial q_k} \eta_k(s) \right]_{\vec{q}=\vec{q}_*} ds + \left[\frac{\partial F}{\partial \dot{q}_k} \eta_k(s) \right]_{t_0}^{t_1} - \int_{t_0}^{t_1} \left[\frac{d}{dt} \frac{\partial F}{\partial \dot{q}_k} \eta_k(s) \right]_{\vec{q}=\vec{q}_*} ds \\ (I.2.8) \quad &= \int_{t_0}^{t_1} \eta_k(s) \left[\frac{\partial F}{\partial q_k} - \frac{d}{dt} \frac{\partial F}{\partial \dot{q}_k} \right]_{\vec{q}=\vec{q}_*} ds. \end{aligned}$$

Nous pouvons donc conclure que les équations d'Euler (I.2.6) sont une *condition suffisante* pour avoir la fonctionnelle stationnaire ; si elles sont vérifiées alors de (I.2.8) suit que $\frac{\partial}{\partial \alpha_k} \mathcal{F}_\eta(0) = 0$ et puisque $\vec{\eta}$ était arbitraire nous pouvons conclure que la fonctionnelle est stationnaire.

Nous allons maintenant montrer que cette condition est aussi nécessaire, c'est-à-dire, nous supposons que \vec{q}_* est un point où la fonctionnelle est stationnaire, et nous allons montrer que \vec{q}_* vérifie les équations d'Euler.

Par hypothèse nous avons (le calcul fait jusqu'ici est toujours valable) :

$$(I.2.9) \quad \frac{\partial}{\partial \alpha_k} \mathcal{F}_\eta(0) = \int_{t_0}^{t_1} \eta_k(s) \left[\frac{\partial F}{\partial q_k} - \frac{d}{dt} \frac{\partial F}{\partial \dot{q}_k} \right]_{\vec{q}=\vec{q}_*} ds = 0 \quad \forall \vec{\eta} \in H \text{ et } k = 1, \dots, m.$$

Notons la fonction à intégrer par $\Phi_k(s) = \eta_k(s) \left[\frac{\partial F}{\partial q_k} - \frac{d}{dt} \frac{\partial F}{\partial \dot{q}_k} \right]$, et observons que :

- (1) cette fonction est *continue*, c'est d'ailleurs pour obtenir cette condition que nous avons demandé des perturbations \mathcal{C}^2 ;
- (2) les fonctions $\eta_k(s)$ sont arbitraires dans H .

Mais alors s'il existait un \bar{t} dans (t_0, t_1) tel que $\Phi_k(\bar{t}) > 0$ pour un certain k (le cas $\Phi_k(\bar{t}) < 0$ est similaire), grâce à la continuité de Φ_k il y aurait un voisinage de \bar{t} , appelons-le \bar{I} , où cette fonction garderait le signe positif. Puisque les fonctions η_k sont arbitraires nous pourrions en choisir une avec signe constant et avec support contenu dans \bar{I} , mais alors l'intégrale de $\Phi_k(s)$ pour ce choix de η_k serait différente de zéro, ce qui contredit l'hypothèse (I.2.9).

On doit donc en conclure que $\Phi_k(s) \equiv 0$ sur (t_0, t_1) pour tout $k = 1, \dots, m$. Mais encore une fois car les fonctions η_k sont arbitraires nous pouvons conclure que :

$$\frac{d}{dt} \frac{\partial F}{\partial \dot{q}_k} - \frac{\partial F}{\partial q_k} = 0 \quad \forall k = 1, \dots, m,$$

c'est-à-dire les équations d'Euler sont vérifiées. \square

Remarquons que l'analogie entre les équations d'Euler (I.2.6) et les équations de Lagrange (I.1.2) est frappante. Observons aussi que si la fonction F ne dépend pas explicitement du temps alors il est de facile vérifier que les équations (I.2.6) admettent comme intégrale première la fonction

$$G = \sum_{k=1}^m \frac{\partial F}{\partial \dot{q}_k} \dot{q}_k - F.$$

En effet :

$$\begin{aligned} \frac{d}{dt} \left[\sum_{k=1}^m \frac{\partial F}{\partial \dot{q}_k} \dot{q}_k - F \right] &= \sum_{k=1}^m \left[\dot{q}_k \frac{d}{dt} \frac{\partial F}{\partial \dot{q}_k} + \frac{\partial F}{\partial \dot{q}_k} \ddot{q}_k \right] - \sum_{k=1}^m \left[\dot{q}_k \frac{\partial F}{\partial q_k} + \ddot{q}_k \frac{\partial F}{\partial \dot{q}_k} \right] \\ &= \sum_{k=1}^m \dot{q}_k \left[\frac{d}{dt} \frac{\partial F}{\partial \dot{q}_k} - \frac{\partial F}{\partial q_k} \right] \equiv 0, \end{aligned} \quad (\text{I.1.2})$$

donc G est constante sur les solutions de (I.1.2).

2.2. Le plus court chemin entre deux points. Montrons qu'avec la méthode du calcul des variations nous pouvons facilement résoudre certains problèmes d'optimisation. Par exemple nous savons depuis la géométrie élémentaire que le *plus court chemin* entre deux points dans \mathbb{R}^2 est un segment de droite.

Fixons deux points P_0 et P_1 dans le plan et considérons toutes les courbes régulières (de classe \mathcal{C}^2) données sous forme paramétrisée $\gamma(t) = (t, f(t))$, telle que $\gamma(t_0) = P_0$ et $\gamma(t_1) = P_1$; pour simplifier la géométrie nous pouvons supposer sans perte de généralité (rotation et déplacement rigide de l'origine) que $f(t_0) = f(t_1) = 0$. Introduisons la fonctionnelle suivante *distance Euclidienne* :

$$(I.2.10) \quad \mathcal{F}(\gamma, \dot{\gamma}) := \int_{t_0}^{t_1} \sqrt{1 + (f'(s))^2} ds.$$

Grâce au théorème précédent nous devons résoudre les équations d'Euler pour la fonction $F(\gamma, \dot{\gamma}) = \sqrt{1 + (f')^2}$:

$$\frac{d}{dt} \frac{\partial F}{\partial f'} - \frac{\partial F}{\partial f} = 0,$$

puisque F ne dépend pas explicitement de f nous avons :

$$\text{constante} = \frac{\partial F}{\partial f'} = \frac{f'}{\sqrt{1 + (f')^2}}.$$

Un simple calcul nous donne :

$$(f')^2 = \frac{\text{constante}^2}{1 - \text{constante}^2},$$

donc nous devons chercher une fonction $f \in \mathcal{C}^2$ avec dérivée constante sur l'intervalle (t_0, t_1) et qui vaut zéro aux extrémités de cet intervalle. La seule possibilité est donc que cette constante soit égale à zéro et donc $f(s) \equiv 0$. Donc la courbe γ qui vérifie les équations d'Euler et qui minimise la fonctionnelle distance Euclidienne est $\gamma(s) = (s, 0)$, c'est-à-dire le segment qui relie les points P_0 et P_1 .

Ce problème de recherche de la courbe qui réalise la distance minimale entre deux points peut être généralisé à une surface quelconque ou une variété de Riemann générique, il faudra définir proprement la nouvelle fonctionnelle distance (*distance*

riemanniennes) et la courbe solution sera alors appelée *géodésique*. Il faut cependant remarquer que les géodésiques sont seulement des points stationnaires de la fonctionnelle distance, on peut bien trouver des cas où elles sont des maxima ; par exemple sur une sphère les géodésiques sont les (arc de) méridiens, pour tout couple de points, non opposés sur un diamètre, alors il y a un seul méridien qui les réunit, mais deux géodésiques : une de longueur minimale et l'autre maximale (voir Fig. 4).

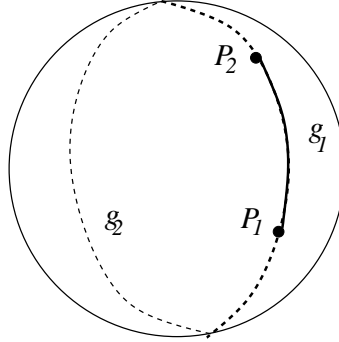


FIG. 4

2.3. Principe variationnel d'Hamilton : formulation Lagrangienne.

Nous avons vu dans le paragraphe précédent l'équivalence entre les solutions des équations d'Euler et le problème de minimiser une certaine fonctionnelle, nous avons aussi vu que pour une particule libre (voir exemple 2.1) le mouvement naturel était celui qui minimisait l'énergie cinétique *moyenne* sur la durée du mouvement. Ces réflexions et l'analogie entre les équations d'Euler et celles d'Euler–Lagrange nous poussent à définir le principe suivant de la mécanique classique :

THÉOREME 2.3 (Principe d'Hamilton). *Le mouvement naturel est caractérisé par la propriété de rendre stationnaire l'Action Hamiltonienne :*

$$(I.2.11) \quad A(\vec{q}, \dot{\vec{q}}) := \int_{t_0}^{t_1} \mathcal{L}(\vec{q}(s), \dot{\vec{q}}(s), s) ds,$$

dans la classe des mouvements variés synchrones ¹ qui préservent les configurations initiales et finales.

Nous l'avons appelé principe car on pourrait le postuler pour vrai et en faire découler toute la mécanique lagrangienne, équations d'Euler–Lagrange comprises.

La démonstration suit exactement celle du Théorème d'Euler pour la fonctionnelle action.

Puisque la fonction lagrangienne est $\mathcal{L} = T - V$, le principe d'Hamilton dit que le mouvement naturel est celui, parmi toutes les courbes possibles, qui rend stationnaire la moyenne temporelle de la différence entre l'énergie cinétique et l'énergie potentielle.

¹Il s'agit de l'ensemble \mathcal{Q} défini précédemment ; on les appelle synchrones car les perturbations sont toutes faites en même temps et la variable temporelle ne varie pas.

CHAPITRE II

Introduction à la mécanique hamiltonienne.

Nous avons vu que l'introduction des coordonnées généralisées permet de réduire le nombre d'équations à résoudre, mais elles restent du deuxième ordre ; une des motivations qui poussent à introduire un nouveau formalisme, le *formalisme Hamiltonien*, est la volonté d'avoir des équations “plus simples”, notamment des équations du premier ordre.

Il s'agit donc de déterminer un changement de coordonnées qui transforme les équations du mouvement dans une forme plus simple, dite *canonique*, qui permettra de mettre en évidence beaucoup de propriétés du système. Pour faire cela nous allons introduire et utiliser une application connue sous le nom de *transformation de Legendre*.

1. La transformation de Legendre

Considérons une fonction réelle d'une variable réelle définie dans un intervalle (a, b) , $-\infty \leq a < b \leq +\infty$, qui soit \mathcal{C}^2 dans son domaine de définition. Supposons aussi que la fonction soit *convexe*, c'est-à-dire :

$$f''(x) > 0 \quad \forall x \in (a, b) .$$

Nous définissons la *transformation de Legendre* de la fonction $f(x)$ la fonction $g(p)$ définie, pour p variable dans un certain intervalle, par :

$$(II.1.1) \quad g(p) := pw(p) - f(w(p)) ,$$

où $w(p)$ est le seul point qui vérifie :

$$(II.1.2) \quad f'(w(p)) = p .$$

Puisque $f'' > 0$, la fonction f' est strictement croissante et donc si l'équation (II.1.2) a une solution, cette solution est unique, de plus elle dépend continûment de p .

Géométriquement la fonction $w(p)$ peut être vue comme l'abscisse du point où le graphe de la fonction $f(x)$ est tangent à la droite de pente p (voir Fig. 1 à gauche). Une deuxième façon de définir la transformation de Legendre est donnée par la remarque suivante.

REMARQUE 1.1. Dans le plan (x, y) on trace le graphe de $y = f(x)$, on se donne un nombre réel p et on considère la droite $y = px$. Soit $x(p)$ le point tel que la distance verticale entre la courbe $y = f(x)$ et la droite $y = px$ est maximale. Considérons la fonction $F(x, p) = px - f(x)$, alors pour p fixé, cette fonction a un extremum en $x(p)$ (voir Fig. 1 à droite), nous définissons donc $g(p) = px(p) - f(x(p), p)$ (distance sur la verticale).

Le point $x(p)$ est déterminé par la condition $\partial F / \partial x = 0$, c'est-à-dire $f'(x(p)) = p$; à cause de la convexité de f en ce point, s'il existe, il est unique. On voit aussi l'équivalence de cette définition avec la précédente.

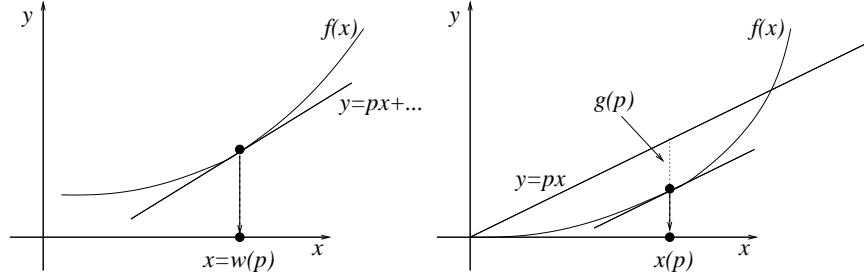


FIG. 1

Le résultat suivant permet d'affirmer que la transformation de Legendre admet à son tour une transformation de Legendre (la transformation de Legendre est *inversible* et *involution*).

THÉOREME 1.2. *La fonction $g(p)$, transformée de Legendre d'une fonction $f \in \mathcal{C}^2$, convexe, admet une transformation de Legendre qui coïncide avec la fonction f du départ.*

DÉMONSTRATION. Nous devons vérifier d'abord que g est régulière et convexe. Quelle que soit la définition utilisée le théorème de la fonction implicite assure que $x(p)$ est une fonction autant régulière que f et donc g aussi. Le calcul de la dérivée seconde de g est trivial ; tout d'abord la dérivée première :

(II.1.3)

$$g'(p) = w(p) + pw'(p) - f'(w(p))w'(p) = w(p) + w'(p)(p - f'(w(p))) = w(p),$$

donc

$$g''(p) = w'(p).$$

En utilisant le théorème de la dérivée de la fonction inverse, nous trouvons :

$$(II.1.4) \quad g''(p) = w'(p) = \frac{1}{f''(p)} > 0.$$

La fonction g est donc convexe et nous pouvons en faire la transformation de Legendre ; nous cherchons un point $p = p(w)$ tel que :

$$g'(p(w)) = w,$$

mais alors (II.1.3) nous dit que $p(w)$ est la fonction inverse de $w = w(p)$. La transformée de Legendre de g est donc

$$h(w) = wp(w) - g(p(w)),$$

et en utilisant la définition de g nous trouvons :

$$h(w) = wp(w) - g(p(w)) = wp(w) - (p(w)w(p(w)) - f(w(p(w)))) = wp(w) - p(w)w + f(w) = f(w),$$

où nous avons utilisé la propriété : $p(w)$ est la fonction inverse de $w(p)$. \square

EXEMPLE 1.3. *Calculons la transformation de Legendre de $f(w) = w^2$.*

Soit p un réel nous cherchons un point $w(p)$ tel que $f'(w(p)) = p$:

$$p = f'(w(p)) = 2w(p),$$

donc $w(p) = p/2$ et finalement

$$g(p) = pw(p) - f(w(p)) = p\frac{p}{2} - \left(\frac{p}{2}\right)^2 = \frac{p^2}{4}.$$

EXEMPLE 1.4 (Inégalité de Young). Calculons la transformation de Legendre de $f(w) = \frac{w^\alpha}{\alpha}$, pour $\alpha > 1$.

Soit p un réel nous cherchons un point $w(p)$ tel que $f'(w(p)) = p$:

$$p = f'(w(p)) = (w(p))^{\alpha-1},$$

donc

$$w(p) = p^{\frac{1}{\alpha-1}},$$

et finalement

$$g(p) = pw(p) - f(w(p)) = pp^{\frac{1}{\alpha-1}} - \frac{1}{\alpha}p^{\frac{\alpha}{\alpha-1}} = p^{\frac{\alpha}{\alpha-1}} \frac{\alpha-1}{\alpha} := \frac{p^\beta}{\beta}.$$

C'est-à-dire la transformée de Legendre de $f(w) = \frac{w^\alpha}{\alpha}$ est la fonction $g(p) = \frac{p^\beta}{\beta}$ où le paramètre β vérifie :

$$\frac{1}{\alpha} + \frac{1}{\beta} = 1.$$

Cet exemple nous permet de démontrer l'inégalité de Young (cas général de l'inégalité de Cauchy) : pour tout α et β strictement plus grands que 1 et pour tout x et p positifs, alors :

$$px \leq \frac{x^\alpha}{\alpha} + \frac{p^\beta}{\beta}.$$

Pour cela considérons à nouveau la fonction $F(x, p) = px - f(x)$; ses dérivées première et seconde, par rapport à x avec p fixé, sont respectivement :

$$\frac{\partial F}{\partial x} = p - f'(x) \text{ et } \frac{\partial^2 F}{\partial x^2} = -f''(x),$$

donc la fonction F atteint un maximum pour $x = x(p)$ (à p fixé) et donc

$$F(x, p) \leq F(x(p), p) = g(p),$$

c'est-à-dire

$$px - f(x) \leq g(p) \rightarrow px \leq f(x) + g(p).$$

Mais alors ça suffit d'écrire la relation précédente pour les fonctions $f(x) = x^\alpha/\alpha$ et sa transformée de Legendre $g(p) = p^\beta/\beta$.

REMARQUE 1.5 (Cas vectoriel). Les considérations précédentes peuvent être généralisées au cas vectoriel. Soit $f : \mathbb{R}^m \rightarrow \mathbb{R}$ une fonction C^2 (dérivées partielles secondes continues) telle que la matrice Hessienne $\frac{\partial^2 f}{\partial w_j \partial w_j}$ soit définie positive. Alors nous pouvons inverser le système :

$$\frac{\partial f}{\partial w_i} = p_i \quad i = 1, \dots, m,$$

où p_i sont des réels, et ainsi définir la fonction $\vec{w} = \vec{w}(\vec{p})$. La transformation de Legendre de f sera alors :

$$g(\vec{p}) := \vec{p} \cdot \vec{w}(\vec{p}) - f(\vec{w}(\vec{p})).$$

REMARQUE 1.6. La transformation de Legendre peut être généralisée au cas des fonctions qui ne sont pas C^2 . Soit donc $f : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe et semicontinue inférieurement, c'est-à-dire pour tout $x \in \mathbb{R}$ on a $\liminf_{y \rightarrow x} f(y) = f(x)$, alors sa transformée de Legendre est la fonction g définie par :

$$g(p) = \sup_{x \in \mathbb{R}} (xp - f(x)) .$$

On peut vérifier que g ainsi définie est convexe, semicontinue inférieurement et sa transformée de Legendre est f .

2. La fonction d'Hamilton

Nous avons vu que l'énergie cinétique est une forme quadratique définie dans les vecteurs $\dot{\vec{q}}$, donc la fonction de Lagrange $\mathcal{L}(\vec{q}, \dot{\vec{q}}, t)$ admet une transformation de Legendre (vectorielle) par rapport à $\dot{\vec{q}}$ pour tout couple (\vec{q}, t) fixé.

Considérons donc le système :

$$(II.2.1) \quad \frac{\partial \mathcal{L}}{\partial \dot{q}_j} = p_j \quad j = 1, \dots, m ,$$

grâce à la condition sur l'énergie cinétique ce système peut être résolu en fonction de $\dot{\vec{q}} = \dot{\vec{q}}(\vec{p}, \vec{q}, t)$. Mais alors nous pouvons définir la fonction d'Hamilton comme la transformation de Legendre de la Lagrangienne :

$$(II.2.2) \quad \mathcal{H}(\vec{p}, \vec{q}, t) := \vec{p} \cdot \dot{\vec{q}} - \mathcal{L}(\vec{q}, \dot{\vec{q}}, t) ,$$

où les vecteurs $\dot{\vec{q}}$ doivent être lus comme solutions du système (II.2.1).

Les variables (\vec{p}, \vec{q}) sont appelées *variables canoniques*, les \vec{p} sont appelés *moments cinétiques*.

EXEMPLE 2.1. Considérons un point de masse m , libre (aucune force n'agit sur ce point) ; la fonction de Lagrange du système est :

$$\mathcal{L}(\vec{q}, \dot{\vec{q}}, t) = \frac{m\dot{\vec{q}}^2}{2} .$$

Considérons le système (II.2.1) :

$$p_j = \frac{\partial \mathcal{L}}{\partial \dot{q}_j} = m\dot{q}_j ,$$

donc sa transformation de Legendre est :

$$\mathcal{H}(\vec{p}, \vec{q}, t) = \vec{p} \cdot \dot{\vec{q}} - \mathcal{L}(\vec{q}, \dot{\vec{q}}, t) = \frac{\vec{p}^2}{m} - \frac{m}{2} \frac{\vec{p}^2}{m^2} = \frac{\vec{p}^2}{2m} .$$

EXEMPLE 2.2. Considérons maintenant un point de masse m soumis à une force qui dérive d'une énergie potentielle V ; la fonction de Lagrange du système est :

$$\mathcal{L}(\vec{q}, \dot{\vec{q}}, t) = \frac{m\dot{\vec{q}}^2}{2} - V(\vec{q}) .$$

Puisque l'énergie potentielle ne dépend pas de $\dot{\vec{q}}$ le système (II.2.1) admet la même solution et donc la transformation de Legendre nous donne :

$$\mathcal{H}(\vec{p}, \vec{q}, t) = \vec{p} \cdot \dot{\vec{q}} - \mathcal{L}(\vec{q}, \dot{\vec{q}}, t) = \frac{\vec{p}^2}{m} - \frac{m}{2} \frac{\vec{p}^2}{m^2} + V(\vec{q}) = \frac{\vec{p}^2}{2m} + V(\vec{q}) .$$

Donc dans un système avec forces conservatives (c'est-à-dire qui dérivent d'une énergie potentielle), sans contraintes dépendantes du temps, l'Hamiltonien représente l'énergie mécanique totale du système.

3. Les équations d'Hamilton

Montrons maintenant que les nouvelles coordonnées (\vec{p}, \vec{q}) ont bien la propriété requise de "baisser" l'ordre des équations différentielles à résoudre pour déterminer la dynamique.

Nous avons vu que \mathcal{H} est la transformée de Legendre de \mathcal{L} , et nous savons aussi que (Théorème 1.2) \mathcal{L} est alors la transformée de Legendre de \mathcal{H} et donc :

$$(II.3.1) \quad \frac{\partial \mathcal{H}}{\partial p_j} = \dot{q}_j \quad \forall j = 1, \dots, m,$$

qui nous donne la variation de \vec{q} en fonction de la fonction d'Hamilton.

Maintenant nous devons chercher la variation de \vec{p} , pour faire cela nous allons calculer grâce à la définition de la transformation de Legendre en supposant \vec{p} et \vec{q} comme variables indépendantes (donc $\dot{\vec{q}}$ dépend de ces variables) :

$$(II.3.2) \quad \frac{\partial \mathcal{H}}{\partial q_j} = \sum_{i=1}^m p_i \frac{\partial \dot{q}_i}{\partial q_j} - \frac{\partial \mathcal{L}}{\partial q_j} - \sum_{i=1}^m \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{\partial \dot{q}_i}{\partial q_j} = \sum_{i=1}^m p_i \frac{\partial \dot{q}_i}{\partial q_j} - \frac{\partial \mathcal{L}}{\partial q_j} - \sum_{i=1}^m p_i \frac{\partial \dot{q}_i}{\partial q_j} = -\frac{\partial \mathcal{L}}{\partial q_j},$$

où nous avons utilisé $\frac{\partial \mathcal{L}}{\partial \dot{q}_i} = p_i$. Mais alors en utilisant les équations d'Euler-Lagrange, réécrites en termes de \vec{p} et \vec{q} :

$$0 = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \frac{\partial \mathcal{L}}{\partial q_j} = \frac{d}{dt} p_j - \frac{\partial \mathcal{L}}{\partial q_j},$$

donc nous pouvons conclure :

$$(II.3.3) \quad \frac{\partial \mathcal{H}}{\partial q_j} = -\frac{\partial \mathcal{L}}{\partial q_j} = -\dot{p}_j \quad \forall j = 1, \dots, m.$$

Nous avons finalement déterminé les équations différentielles qui déterminent la variation des coordonnées canoniques, ces équations sont appelées *Équations d'Hamilton* :

$$(II.3.4) \quad \begin{cases} \dot{p}_j &= -\frac{\partial \mathcal{H}}{\partial q_j} \\ \dot{q}_j &= \frac{\partial \mathcal{H}}{\partial p_j} \end{cases} \quad \forall j = 1, \dots, m.$$

REMARQUE 3.1. Ces équations sont bien du premier ordre, mais il s'agit d'un système de $2m$ équations, contre un système de m équations du second ordre pour Euler-Lagrange. D'une certaine façon on n'aurait pas pu demander plus que ça pour une sorte de "principe de conservation des difficultés".

Observons aussi que l'Hamiltonien a la même régularité que le Lagrangien et donc les hypothèses d'existence et d'unicité des équations d'Hamilton sont vérifiées.

Dans cette construction nous avons supposé le temps jouant le rôle d'un paramètre, c'est-à-dire qu'il y avait dépendance en temps mais on le considérait comme fixe, nous pouvons maintenant nous demander comment la fonction d'Hamilton évolue dans le temps :

PROPOSITION 3.2. Si $\vec{q}(t)$ et $\vec{p}(t)$ sont les solutions de (II.3.4) alors :

$$(II.3.5) \quad \frac{d}{dt} \mathcal{H}(\vec{p}(t), \vec{q}(t), t) = \frac{\partial \mathcal{H}(\vec{p}(t), \vec{q}(t), t)}{\partial t}.$$

DÉMONSTRATION. La démonstration est immédiate :

$$\begin{aligned} \frac{d}{dt} \mathcal{H}(\vec{p}(t), \vec{q}(t), t) &= \frac{\partial \mathcal{H}(\vec{p}(t), \vec{q}(t), t)}{\partial \vec{p}} \cdot \frac{d\vec{p}}{dt} + \frac{\partial \mathcal{H}(\vec{p}(t), \vec{q}(t), t)}{\partial \vec{q}} \cdot \frac{d\vec{q}}{dt} + \frac{\partial \mathcal{H}(\vec{p}(t), \vec{q}(t), t)}{\partial t} \\ &= \dot{\vec{q}} \cdot \frac{d\vec{p}}{dt} - \dot{\vec{p}} \cdot \frac{d\vec{q}}{dt} + \frac{\partial \mathcal{H}(\vec{p}(t), \vec{q}(t), t)}{\partial t} = \frac{\partial \mathcal{H}(\vec{p}(t), \vec{q}(t), t)}{\partial t}. \end{aligned}$$

□

Ce résultat implique que si l'Hamiltonien ne dépend pas explicitement du temps alors (II.3.5) est une *intégrale première* des équations d'Hamilton, autrement dit la valeur numérique \mathcal{H} est une constante.

4. Théorème de Liouville

Les équations d'Hamilton ont une certaine structure symétrique qui peut être mise en évidence en introduisant un espace à $2m$ dimensions : $\vec{x} = \begin{pmatrix} \vec{p} \\ \vec{q} \end{pmatrix}$ et la matrice $2m \times 2m$:

$$(II.4.1) \quad \mathcal{J} = \begin{pmatrix} \mathbb{O} & -\mathbb{I} \\ \mathbb{I} & \mathbb{O} \end{pmatrix},$$

où \mathbb{I} est la matrice identité $m \times m$ et \mathbb{O} la matrice nulle $m \times m$. Alors les équations (II.3.4) peuvent être réécrites comme :

$$(II.4.2) \quad \dot{\vec{x}} = \mathcal{J} \nabla \mathcal{H},$$

où l'opérateur de dérivation doit être lu comme $\nabla = (\partial_{p_1}, \dots, \partial_{p_m}, \partial_{q_1}, \dots, \partial_{q_m})^T$. Nous avons donc une EDO du premier ordre dans un espace à $2m$ dimensions, et si \mathcal{H} ne dépend pas de façon explicite du temps, alors les trajectoires, i.e. solutions de (II.4.2), appartiennent aux ensembles de niveau $\mathcal{H} = \text{constante}$.

Un champ de vecteur de la forme (II.4.2) a une propriété supplémentaire si \mathcal{H} est assez régulière :

PROPOSITION 4.1. Si \mathcal{H} a des dérivées secondes continues par rapport à \vec{q} et \vec{p} , alors :

$$(II.4.3) \quad \text{div} \vec{X} = 0,$$

où

$$\vec{X} = \mathcal{J} \nabla \mathcal{H}.$$

DÉMONSTRATION. La démonstration découle du calcul suivant (se rappeler que $\vec{x} = (\vec{p}, \vec{q})^T$) :

$$\text{div} \vec{X} = \sum_{j=1}^m \left(\frac{\partial X_j}{\partial p_j} + \frac{\partial X_{m+j}}{\partial q_j} \right),$$

mais alors grâce à (II.4.2) nous avons :

$$\text{div} \vec{X} = \sum_{j=1}^m \left(-\frac{\partial}{\partial p_j} \frac{\partial \mathcal{H}}{\partial q_j} + \frac{\partial}{\partial q_j} \frac{\partial \mathcal{H}}{\partial p_j} \right) = 0,$$

où la dernière égalité est obtenue grâce au Théorème de Schwartz. □

Les équations d'Hamilton ont une propriété encore plus forte car toutes les projections de \vec{X} sur chacun des plans (p_k, q_k) sont aussi de divergence nulle.

Les champs de vecteurs à divergence zéro, aussi appelés *champs solénoïdaux*, ont la remarquable propriété géométrique de *préserver les volumes*.

Les équations d'Hamilton permettent de définir une application dépendante du temps, le *flot Hamiltonien* Φ^t , de l'espace des phases en lui-même ; à tout point $\vec{x}_0 \in \mathbb{R}^{2m}$ et pour tout $t > 0$ on peut lui associer le point $\vec{x}(t)$ obtenu comme solution de (II.4.2) au temps t avec donnée initiale \vec{x}_0 :

$$\vec{x}(t) = \Phi(t, \vec{x}_0).$$

On peut voir de façon très simple que cette transformation est inversible à cause de l'inversibilité des équations du mouvement.

Nous pouvons penser maintenant à transformer à l'aide de Φ^t tout un volume de points dans \mathbb{R}^{2m} et nous demander comment la mesure de ce volume varie dans le temps. Le Théorème suivant dû à Liouville nous donne la réponse : la mesure de ce volume ne change pas, la forme au temps t pourra être très différente de la forme initiale, mais pas son volume.

THÉORÈME 4.2 (Liouville : volume espace des phases). *Le flot Hamiltonien préserve les volumes dans l'espace des phases. Cette propriété est vraie pour tout système de la forme :*

$$(II.4.4) \quad \dot{\vec{x}} = \vec{X}(\vec{x}, t) \quad \text{tel que} \quad \text{div} \vec{X} = 0.$$

DÉMONSTRATION. Il faut démontrer que pour tout $t > 0$ (le cas $t < 0$ est complètement analogue à cause de l'inversibilité du flot) l'image au temps t à travers le flot Hamiltonien, $\Omega(t)$, d'un quelconque domaine $\Omega \subset \mathbb{R}^{2m}$, avec une frontière régulière, a la même mesure que Ω .

La première étape consiste donc à démontrer que la mesure d'un domaine varie dans le temps à cause du flot selon la loi :

$$(II.4.5) \quad \frac{d}{dt} |\Omega(t)| = \int_{\Omega(t)} \text{div} \vec{X}(\vec{x}, t) d\vec{x},$$

où nous avons noté par $|\Omega|$ la mesure du domaine Ω . C'est clair qu'une fois la relation (VI.2.5) démontrée le Théorème est prouvé.

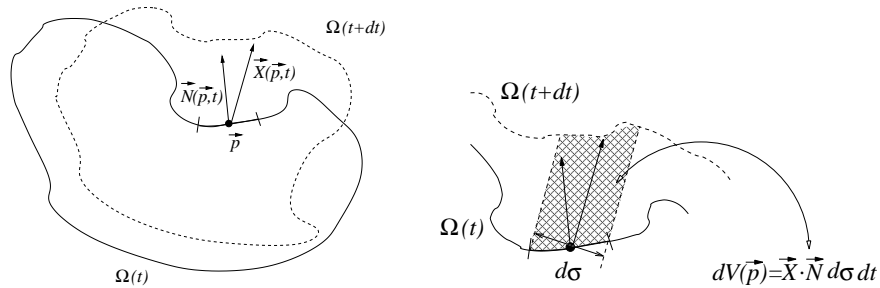


FIG. 2

Dans la Fig. (2) nous avons présenté la géométrie pour le calcul de la variation du volume ; dans la partie de gauche nous avons représenté le domaine Ω à deux instants de temps très proches : t et $t + dt$, nous avons aussi fixé un point \vec{p} sur $\partial\Omega(t)$ où nous avons tracé le champ de vecteur évalué sur ce point, $\vec{X}(\vec{p})$, et le vecteur, $\vec{N}(\vec{p})$, perpendiculaire à l'élément infinitésimal de surface $d\sigma$. Dans la partie de droite nous avons calculé la variation infinitésimale de volume au temps t due au mouvement du point \vec{p} , qui peut être considéré comme solidaire avec la surface $d\sigma$ à cause du champ de vecteur, ce volume est donné en première approximation par un cylindre de base S et de hauteur h , donc $dV(\vec{p}, t) = Sh$. Or la hauteur est donnée par le déplacement du point \vec{p} dans le temps dt , donc $h = \vec{X}(\vec{p})dt$, par contre la base S doit être considérée comme la surface, $d\sigma$ orientée dans la direction du champ de vecteurs, c'est-à-dire $S = d\sigma \vec{N}$, et finalement :

$$(II.4.6) \quad dV(\vec{p}, t) = d\sigma \vec{N} \cdot \vec{X}(\vec{p}) dt.$$

La variation totale de volume en passant de $\Omega(t)$ à $\Omega(t + dt)$ est obtenue en intégrant la contribution (II.4.6) sur tous les points de la frontière :

$$(II.4.7) \quad d|\Omega(t)| = dV(t) = \int_{\partial\Omega(t)} \vec{N} \cdot \vec{X}(\vec{p}) d\sigma dt.$$

Pour exprimer le résultat précédent en terme d'intégrale de volume, nous allons utiliser la *formule de la divergence*¹ :

$$(II.4.8) \quad d|\Omega(t)| = dV(t) = \int_{\Omega(t)} \text{div} \vec{X} d\vec{x} dt,$$

et finalement la (VI.2.5) suit en divisant par dt et en passant à la limite :

$$\frac{d|\Omega(t)|}{dt} = \int_{\Omega(t)} \text{div} \vec{X} d\vec{x}.$$

La conclusion du Théorème est donc triviale, étant donné que le champ a une divergence nulle, cette dernière intégrale est égale à zéro et donc $|\Omega(t)|$ est une fonction constante en t . \square

Ce résultat est très important et il admet plusieurs corollaires, nous allons ici nous concentrer sur deux : Corollaire 4.3 et le Théorème de Poincaré (cfr. Théorème 5.1).

COROLLAIRE 4.3. *Tout point singulier d'un système (II.4.4) ne peut pas être asymptotiquement stable.*

DÉMONSTRATION. Soit \vec{x}_0 un point singulier du système (II.4.4) et supposons qu'il soit asymptotiquement stable. Prenons un sphère centrée sur ce point et de rayon assez petit, tous ses points tendraient vers \vec{x}_0 et donc son volume tendrait à zéro, ce qui contredit le Théorème de Liouville. \square

¹Dans la littérature cette formule peut être appelée avec plusieurs noms, formule de : Gauss, Green, Ostrogradskij, Stokes.

5. Théorème de Poincaré : Récurrence

Le Théorème de Liouville affirme que le flot Hamiltonien ne modifie pas les volumes dans l'espace des phases, par contre la "forme" de ces domaines peut changer dans le temps ; cela a comme conséquence que, dans les hypothèses que le système soit autonome et que la région d'espace des phases accessible soit bornée, le système passera à nouveau, *dans un futur non précisé par des configurations arbitrairement proches de celles de départ*, donc comme si on revenait au début du mouvement. Ce résultat peut paraître contraire à la réalité ; supposons en effet qu'on ouvre un flacon de parfum dans une pièce fermée, alors l'expérience commune est que au bout d'un certain temps le parfum sera évaporé et il sera dispersé dans toute la pièce, et vous pourrez le sentir partout, or le Théorème de Poincaré affirme que si vous attendez *suffisamment* alors une grande partie des particules de parfum rentreront dans le flacon !!

THÉOREME 5.1 (Poincaré, récurrence). *Considérons un système Hamiltonien autonome pour lequel seulement une partie bornée, Ω , d'espace des phases est accessible. Soit $B(0)$ une quelconque sphère contenue dans Ω et soit $B(t)$ son image après un temps t due au flot Hamiltonien. Alors quel que soit $\tau > 0$ il existe $t_0 > \tau$ tel que $B(t_0) \cap B(0) \neq \emptyset$.*

DÉMONSTRATION. Fixons un $\tau > 0$ et considérons la suite de domaines $B_n = B(n\tau)$ pour $n \in \mathbb{N}$. Grâce au Théorème de Liouville ces ensembles ont tous la même mesure. Puisque le système est autonome B_n est obtenu utilisant n fois l'application, M le flot, qui envoie $B(0)$ sur B_1 , c'est-à-dire $B_n = M^{\circ n} B(0)$.

Par hypothèse tous les B_n sont contenus dans Ω qui est borné et donc il doit exister deux entiers n_0 et k tels que :

$$B_{n_0} \cap B_{n_0+k} \neq \emptyset,$$

car autrement l'ensemble $E_N = \cup_{n=1}^N B_n$, contenu dans Ω , aurait une mesure, $N B_0$, de plus en plus grande ce qui contredit l'hypothèse que Ω soit borné.

Considérons maintenant l'ensemble $B_{n_0} \cap B_{n_0+k}$, si $n_0 = 0$ le Théorème est démontré ; supposons donc $n_0 \geq 1$ et parcourons en arrière dans le temps l'image de ces points pour un temps τ , ces points doivent venir respectivement des domaines B_{n_0-1} et B_{n_0+k-1} qui doivent avoir une intersection non vide (utiliser le Théorème de Liouville sur le domaine intersection). Ainsi si nous répétons cette procédure en arrière dans le temps de $n_0\tau$ instants, nous trouvons :

$$B_0 \cap B_k \neq \emptyset.$$

□

6. Méthodes variationnelles en mécanique Hamiltonienne

Nous avons introduit au paragraphe I.2 une formulation différente de la mécanique classique, non plus fondée sur les équations d'Euler–Lagrange mais sur le *principe (variationnel) d'Hamilton* : le mouvement naturel est celui qui rend stationnaire l'action Hamiltonienne dans la classe des mouvements variés synchrones.

L'action Hamiltonienne est construite à partir du lagrangien du système, nous pouvons cependant en passant au formalisme hamiltonien, la réécrire en termes de

l'hamiltonien du système, nous aurons donc :

$$(II.6.1) \quad A(\vec{p}, \vec{q}) := \int_{t_0}^{t_1} [\vec{p} \cdot \dot{\vec{q}} - \mathcal{H}(\vec{p}, \vec{q}, t)] dt,$$

où nous avons sous-entendu de devoir exprimer les vitesses généralisées en termes des nouvelles variables canoniques. $\dot{\vec{q}} = \dot{\vec{q}}(\vec{p}, \vec{q}, t)$. En outre les mouvements variés doivent être définis sur l'espace des phases (\vec{p}, \vec{q}) . On pourrait perturber les \vec{q} et déterminer la variation résultante sur les \vec{p} de façon à préserver la relation $p_k = \frac{\partial \mathcal{L}}{\partial \dot{q}_k}$. Mais il s'avère plus intéressant d'effectuer des variations indépendantes sur les deux ensembles de coordonnées :

$$(II.6.2) \quad \begin{cases} q_k(t) &= q_k^*(t) + \eta_k(t) & k = 1, \dots, m \\ p_k(t) &= p_k^*(t) + \zeta_k(t) & k = 1, \dots, m, \end{cases}$$

avec $\vec{\eta}$ et $\vec{\zeta}$ fonctions $\mathcal{C}^2[t_0, t_1]$, $\vec{\eta}(t_0) = \vec{\eta}(t_1) = 0$ et (\vec{p}^*, \vec{q}^*) est une solution de "repère".

Remarquons que nous n'imposons pas ici la condition aux extrêmes pour les variations des \vec{p} , donc cette classe de variations satisfait toujours aux contraintes d'être aux points P_0 et P_1 aux temps t_0 et t_1 . En général les courbes (II.6.2) ne sont pas des solutions des équations d'Hamilton, par exemple si $m = 1$ pour un point libre nous avons $p = m\dot{q}$ et si $\zeta(t) \neq m\dot{\eta}(t)$ alors il n'y a plus la correspondance $p = \frac{\partial \mathcal{L}}{\partial \dot{q}}$. Donc cette classe de trajectoires est plus grande que celle des mouvements variés synchrones.

Nous pouvons donc démontrer le résultat suivant :

THÉOREME 6.1. *Une condition nécessaire et suffisante pour rendre stationnaire l'action Hamiltonienne (II.6.1) dans la classe des mouvements (II.6.2) est que (\vec{p}^*, \vec{q}^*) soient solutions des équations d'Hamilton.*

DÉMONSTRATION. Si nous écrivons la variation de l'action :

$$(II.6.3) \quad \delta A = \int_{t_0}^{t_1} [\vec{p} \cdot \dot{\vec{q}} - \mathcal{H}(\vec{p}, \vec{q}, t)] dt - \int_{t_0}^{t_1} [\vec{p}^* \cdot \dot{\vec{q}}^* - \mathcal{H}(\vec{p}^*, \vec{q}^*, t)] dt,$$

alors en utilisant (II.6.2) nous pouvons obtenir à l'ordre 1 dans les perturbations $\vec{\zeta}$ et $\vec{\eta}$:

$$(II.6.4) \quad \delta A = \int_{t_0}^{t_1} \left[\vec{\zeta} \cdot \dot{\vec{q}}^* + \vec{p}^* \cdot \dot{\vec{\eta}} - \sum_{k=1}^m \left(\frac{\partial \mathcal{H}(\vec{p}^*, \vec{q}^*)}{\partial p_k} \zeta_k + \frac{\partial \mathcal{H}(\vec{p}^*, \vec{q}^*)}{\partial q_k} \eta_k \right) \right] dt.$$

Une intégration par partie sur le terme avec $\dot{\vec{\eta}}$ en utilisant la condition sur les fonctions $\vec{\eta}$ aux extrêmes d'intégration, permet de réécrire (II.6.4) sous la forme :

$$(II.6.5) \quad \delta A = \int_{t_0}^{t_1} \sum_{k=1}^m \left[\zeta_k \left(\dot{q}_k - \frac{\partial \mathcal{H}(\vec{p}^*, \vec{q}^*)}{\partial p_k} \right) + \eta_k \left(-\dot{p}_k - \frac{\partial \mathcal{H}(\vec{p}^*, \vec{q}^*)}{\partial q_k} \right) \right] dt.$$

Nous pouvons donc conclure que les équations d'Hamilton sont une condition suffisante pour rendre stationnaire l'action hamiltonienne, et de façon similaire à celle utilisée dans le Théorème I.2.2 (utiliser le fait que les perturbations sont arbitraires et la régularité des fonctions) nous pouvons aussi conclure la nécessité des équations d'Hamilton. \square

Il existe plusieurs principes variationnels dans le formalisme hamiltonien, nous terminons cette section en présentant le principe (le plus connu), celui *de l'action stationnaire* ou *principe de Maupertuis*, qui nous permettra d'introduire les perturbation asynchrones.

6.1. Principe de l'action stationnaire ou de Maupertuis. Ce principe est valide pour des systèmes hamiltoniens où la fonction d'Hamilton ne dépend pas de façon explicite du temps. On considère donc l'espace des phases généralisé, c'est-à-dire l'espace \mathbb{R}^{2m+1} des triples (\vec{p}, \vec{q}, t) , et on paramétrise les courbes avec une nouvelle variable u :

$$(II.6.6) \quad \vec{p} = \vec{p}(u) \quad \vec{q} = \vec{q}(u) \quad \text{et} \quad t = t(u).$$

Considérons maintenant un mouvement naturel dans la paramétrisation précédente :

$$(II.6.7) \quad \vec{p}^* = \vec{p}^*(u) \quad \vec{q}^* = \vec{q}^*(u) \quad \text{et} \quad t = t^*(u),$$

où la fonction $t^*(u)$ est définie pour $u_0 \leq u \leq u_1$, elle est \mathcal{C}^2 dans le même intervalle fermé, et telle que :

$$(II.6.8) \quad t^*(u_0) = t_0 \quad t^*(u_1) = t_1 \quad \text{et} \quad \frac{dt^*}{du} \neq 0 \quad \forall u \in [u_0, u_1].$$

Nous pouvons donc introduire des “perturbations”, $\vec{\zeta}, \vec{\eta}$ et τ :

$$(II.6.9) \quad \vec{p}(u) = \vec{p}^*(u) + \vec{\zeta}(u) \quad \vec{q}(u) = \vec{q}^*(u) + \vec{\eta}(u) \quad \text{et} \quad t(u) = t^*(u) + \tau(u),$$

régulières \mathcal{C}^2 et telles que $\vec{\eta}(u_0) = \vec{\eta}(u_1) = 0$.

Cette classe de variation est encore plus ample que celle définie par (II.6.2) car maintenant les perturbations sur les variables canoniques peuvent être faites à des instants de temps différents, *mouvements variés asynchrones*.

Puisque la fonction d'Hamilton ne dépend pas de façon explicite du temps, nous savons que l'énergie du système, c'est-à-dire la valeur numérique de \mathcal{H} , ne change pas sur le mouvement naturel. Or les perturbations (II.6.9) ne sont pas forcément des mouvements solutions des équations d'Hamilton et donc l'énergie pourrait ne pas être préservée par ces perturbations; nous allons donc nous restreindre à la classe des perturbations (II.6.9) *isoénergétique*, c'est-à-dire telle que l'énergie soit constante :

$$(II.6.10) \quad \mathcal{H}(\vec{p}(u), \vec{q}(u)) = \mathcal{H}(\vec{p}^*(u), \vec{q}^*(u)) \quad \forall u \in [u_0, u_1].$$

Nous pouvons finalement énoncer le résultat suivant :

THÉORÈME 6.2. *Si la fonction d'Hamilton ne dépend pas du temps de façon explicite alors la fonctionnelle Action de Maupertuis :*

$$(II.6.11) \quad A_{Maup} := \int_{t(u_0)}^{t(u_1)} \vec{p} \cdot \dot{\vec{q}} dt,$$

est stationnaire sur la classe des mouvements variés asynchrones (II.6.9) isoénergétiques (II.6.10).

DÉMONSTRATION. Réécrivons l'action de Maupertuis en utilisant la forme des perturbation (II.6.9) et en faisant l'intégration sur la variable u :

$$(II.6.12) \quad A_{Maup} = \int_{u_0}^{u_1} \left(\vec{p}^*(u) + \vec{\zeta}(u) \right) \cdot \frac{d}{du} (\vec{q}^*(u) + \vec{\eta}(u)) du;$$

la variation de l'action de Maupertuis au premier ordre dans les perturbations est (nous avons fait une intégration par partie pour "déplacer" la dérivée par rapport à u de $\vec{\eta}$ à \vec{p} et nous avons utilisé la propriété des fonctions $\vec{\eta}$ aux extrêmes d'intégration) :

$$(II.6.13) \quad \delta A_{Maup} = \int_{u_0}^{u_1} \left(\vec{\zeta} \cdot \frac{d\vec{q}^*}{du} - \frac{d\vec{p}^*}{du} \cdot \vec{\eta} \right) du.$$

Les équations d'Hamilton permettent d'écrire :

$$(II.6.14) \quad \frac{d\vec{q}^*}{du} = \frac{d\vec{q}^*}{dt} \frac{dt}{du} = \frac{\partial \mathcal{H}}{\partial \vec{p}} \frac{dt}{du},$$

et de façon similaire :

$$\frac{d\vec{p}^*}{du} = - \frac{\partial \mathcal{H}}{\partial \vec{q}} \frac{dt}{du},$$

d'où nous pouvons calculer la variation (au premier ordre) de la fonction d'Hamilton (car elle ne dépend pas du temps) :

$$(II.6.15) \quad \begin{aligned} \delta \mathcal{H} &:= \mathcal{H}(\vec{\zeta} + \vec{p}^*, \vec{\eta} + \vec{q}^*) - \mathcal{H}(\vec{p}^*, \vec{q}^*) = \frac{\partial \mathcal{H}}{\partial \vec{p}} \cdot \vec{\zeta} + \frac{\partial \mathcal{H}}{\partial \vec{q}} \cdot \vec{\eta} \\ &= \left(\frac{d\vec{q}^*}{du} \cdot \vec{\zeta} - \frac{d\vec{p}^*}{du} \cdot \vec{\eta} \right) \frac{du}{dt}. \end{aligned}$$

Le calcul de (II.6.15) permet de récrire la variation de l'action de Maupertuis (II.6.13) comme :

$$(II.6.16) \quad \delta A_{Maup} = \int_{u_0}^{u_1} \frac{dt}{du} \delta \mathcal{H} du,$$

mais par hypothèse les perturbations sont isoénergétiques, donc $\delta \mathcal{H} = 0$, donc nous pouvons conclure que l'action est stationnaire dans cette classe de mouvements variés asynchrones isoénergétiques. \square

CHAPITRE III

Formalisme canonique.

1. Introduction

Dans ce chapitre nous allons présenter certains aspects plus géométriques propres à la formulation hamiltonienne de la mécanique classique. D'après certains auteurs la mécanique hamiltonienne est une forme de géométrie (voir Fig. 1 où nous reportons la page d'introduction du manuel *méthodes analytiques de mécanique classique* de V. Arnol'd).

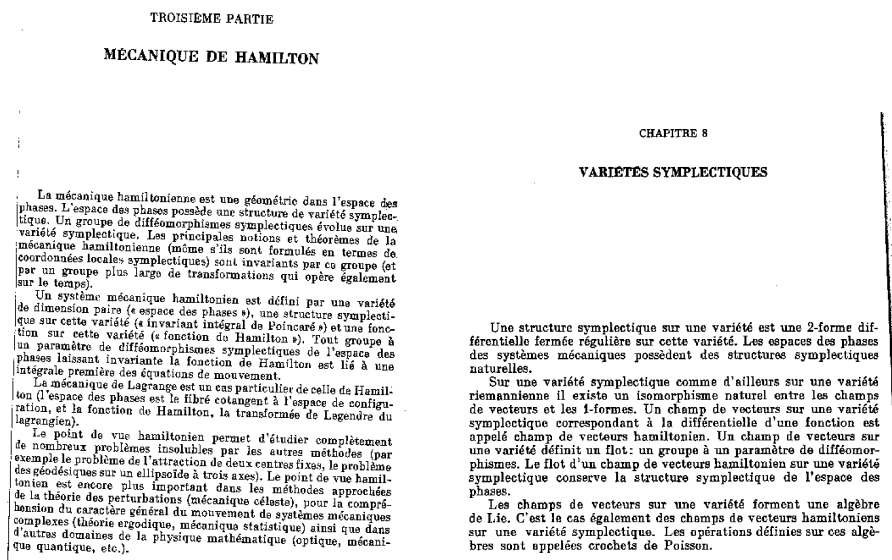


FIG. 1

Ce point de vue est fort général et il s'avère très utile quand nous sommes intéressés à des questions globales (donc liées à la géométrie du problème), par contre si nous travaillons en coordonnées locales les méthodes introduites dans le chapitre précédent et dans les suivants seront suffisantes. Nous avons aussi inséré la section 5 dont le but est de proposer aux étudiants un bref rappel des notions de : vecteur, forme différentielle sur une variété, normalement vues dans le cours de Géométrie Différentielle.

Il s'agit ici d'un problème similaire à celui qu'on voit avec les surfaces ou les variétés : de manière générale il faut plusieurs systèmes de coordonnées locales

pour pouvoir “couvrir” toute la surface ou variété, de plus ces systèmes de coordonnées doivent se “raccorder convenablement” sur les domaines communs ; de plus on doit pouvoir exprimer certaines propriétés “intrinsèques” de la surface ou variété de façon indépendante des coordonnées choisies. En mécanique hamiltonienne la situation est analogue, par exemple on pourrait se demander si la conservation de l’hamiltonien pour un système avec forces conservatives et indépendant du temps, vue dans le chapitre précédent dans des coordonnées locales particulières, (\vec{p}, \vec{q}) , est une propriété propre du système, et donc dite *géométrique*, ou dépend du choix des coordonnées et dans ce cas comment cette valeur change-t-elle en changeant de coordonnées.

La géométrie symplectique peut être vue comme une possibilité d’écrire la mécanique hamiltonienne de façon intrinsèque, c’est-à-dire indépendamment des coordonnées choisies.

2. Structure symplectique de l’espace des phases.

Dans la section II.4 nous avons introduit la matrice réelle $2m \times 2m$, \mathcal{J} :

$$(III.2.1) \quad \mathcal{J} = \begin{pmatrix} \mathbb{O} & -\mathbb{I} \\ \mathbb{I} & \mathbb{O} \end{pmatrix},$$

on vérifie aisément que la matrice \mathcal{J} est *orthogonale* et *antisymétrique* :

$$(III.2.2) \quad \mathcal{J}^{-1} = (\mathcal{J})^T = -\mathcal{J} \quad \text{et} \quad \mathcal{J}^2 = -\mathbb{I}.$$

Nous avons aussi vu que l’introduction du vecteur $\vec{x} = (\vec{p}, \vec{q})^T$, permet de récrire les équations d’Hamilton sous la forme (compacte) :

$$(III.2.3) \quad \dot{\vec{x}} = \mathcal{J} \nabla_{\vec{x}} \mathcal{H}(\vec{x}, t).$$

EXEMPLE 2.1. *Considérons un système Hamiltonien linéaire, c’est-à-dire*

$$(III.2.4) \quad \mathcal{H}(\vec{x}) = \frac{1}{2} (\vec{x})^T S \vec{x} = \sum_{i,j=1}^{2m} x_i S_{ij} x_j,$$

où S est une matrice réelle symétrique. Les équations d’Hamilton sont donc :

$$\dot{\vec{x}} = \mathcal{J} S \vec{x},$$

dont la solution explicite est donnée par :

$$(III.2.5) \quad \vec{x}(t) = e^{(t-t_0)B} \vec{x}(t_0),$$

où par simplicité nous avons introduit la matrice $B = \mathcal{J}S$. Cette matrice vérifie une importante propriété :

$$(B)^T \mathcal{J} + \mathcal{J} B = 0,$$

en effet nous avons :

$$(B)^T \mathcal{J} + \mathcal{J} B = (S)^T (\mathcal{J})^T \mathcal{J} + \mathcal{J} S = S(-\mathcal{J})\mathcal{J} - \mathbb{I}S = -S + S = 0.$$

Cette propriété joue un rôle fondamental dans la théorie, c’est pour cela qu’on donne la définition suivante :

DÉFINITION 2.2 (Matrice hamiltonienne). *Toute matrice réelle, carrée, avec un nombre paire de lignes, A , est dite matrice hamiltonienne si :*

$$(III.2.6) \quad (A)^T \mathcal{J} + \mathcal{J} A = 0.$$

Les matrices hamiltoniennes sont caractérisées par la propriété suivante :

THÉOREME 2.3. *Les propriétés suivantes sont équivalentes :*

- (1) *A est hamiltonienne ;*
- (2) *il existe une matrice symétrique S telle que $A = \mathcal{J}S$;*
- (3) *la matrice $\mathcal{J}A$ est symétrique.*

En outre si A et B sont deux matrices hamiltoniennes alors :

$$(A)^T, \quad \lambda A \quad \forall \lambda \in \mathbb{R}, \quad A \pm B \quad \text{et} \quad [A, B] := AB - BA,$$

sont toutes des matrices hamiltoniennes.

DÉMONSTRATION. L'équivalence entre (2) et (3) est banale car, s'il existe S symétrique telle que $A = \mathcal{J}S$ alors, $\mathcal{J}A = \mathcal{J}\mathcal{J}S = -S$ qui est symétrique. Réciproquement, $A = -\mathcal{J}(\mathcal{J}A)$ et si $\mathcal{J}A$ est symétrique, alors nous avons déterminé $S = -\mathcal{J}A$, symétrique, telle que $A = \mathcal{J}S$.

Par la définition de matrice hamiltonienne nous avons :

$$\mathcal{J}A = -(A)^T \mathcal{J} = (A)^T (\mathcal{J})^T = (\mathcal{J}A)^T,$$

où nous avons utilisé la propriété de la matrice \mathcal{J} , $(\mathcal{J})^T = -\mathcal{J}$, la règle de transposition d'un produit. Donc la matrice $\mathcal{J}A$ est symétrique : (1) implique (3). Et de façon triviale (3) implique (1), si $\mathcal{J}A$ est symétrique alors :

$$\mathcal{J}A = (\mathcal{J}A)^T = (A)^T (\mathcal{J})^T = -(A)^T \mathcal{J} \Rightarrow \mathcal{J}A + (A)^T \mathcal{J} = 0.$$

Les trois premières affirmations sont triviales, nous nous contentons de montrer la dernière. Grâce au point (2) il existe S et R matrices réelles symétriques telles que : $A = \mathcal{J}S$ et $B = \mathcal{J}R$, alors le commutateur est :

$$[A, B] = AB - BA = \mathcal{J}S\mathcal{J}R - \mathcal{J}R\mathcal{J}S = \mathcal{J}(S\mathcal{J}R - R\mathcal{J}S),$$

mais

$$(S\mathcal{J}R - R\mathcal{J}S)^T = (R)^T (\mathcal{J})^T (S)^T - (S)^T (\mathcal{J})^T (R)^T = -R\mathcal{J}S + S\mathcal{J}R = -(R\mathcal{J}S - S\mathcal{J}R),$$

où nous avons utilisé le fait que S et R sont symétriques et la propriété de la transposition de \mathcal{J} , donc :

$$\begin{aligned} \mathcal{J}[A, B] &= \mathcal{J}\mathcal{J}(S\mathcal{J}R - R\mathcal{J}S) = -(S\mathcal{J}R - R\mathcal{J}S) = -(S\mathcal{J}R - R\mathcal{J}S)^T = (S\mathcal{J}R - R\mathcal{J}S)^T \mathcal{J} \\ &= -(S\mathcal{J}R - R\mathcal{J}S)^T (\mathcal{J})^T \mathcal{J} = -(\mathcal{J}(S\mathcal{J}R - R\mathcal{J}S))^T \mathcal{J} \\ &= -([A, B])^T \mathcal{J}, \end{aligned}$$

c'est-à-dire $[A, B]$ est hamiltonienne. □

Ce théorème montre que les matrices hamiltoniennes forment un (sous)groupe des matrices réelles carrées, par rapport à l'addition, ce groupe est appelé : $sp(m, \mathbb{R})$. En outre, comme le commutateur peut être vu comme un produit, et alors on l'appelle *produit de Lie*, ce groupe a une structure d'*algèbre de Lie*, c'est pour cela qu'on l'indique avec une lettre minuscule, comme on le fait normalement avec l'algèbre associée à un groupe donné. Le groupe en question est le groupe des *matrices symplectiques* :

DÉFINITION 2.4 (Matrices Symplectiques). *Une matrice réelle, carrée est dite symplectique si :*

$$(III.2.7) \quad (A)^T \mathcal{J}A = \mathcal{J}.$$

Nous avons la caractérisation suivante :

THÉOREME 2.5. *Les matrices symplectiques $2m \times 2m$ forment un groupe, dénoté par $Sp(m, \mathbb{R})$, par rapport au produit de matrices. La transposée d'une matrice symplectique est symplectique.*

DÉMONSTRATION. L'identité du groupe est la matrice identité $2m \times 2m$, car elle est trivialement symplectique, de plus si A vérifie (III.2.7) elle n'est pas singulière et donc elle est inversible, car :

$$\det \mathcal{J} = \det((A)^T \mathcal{J} A) = \det(A)^T \det \mathcal{J} \det A \Rightarrow (\det A)^2 = 1.$$

De plus son inverse vérifie $A^{-1} = \mathcal{J}^{-1}(A)^T \mathcal{J} = -\mathcal{J}(A)^T \mathcal{J}$, donc :

$$(A^{-1})^T \mathcal{J} A^{-1} = -(-\mathcal{J})A(\mathcal{J})^T \mathcal{J} A^{-1} = -(\mathcal{J})^T A \mathcal{J}^{-1} \mathcal{J} A^{-1} = \mathcal{J} A A^{-1} = \mathcal{J},$$

c'est-à-dire que A^{-1} est aussi symplectique.

Soit maintenant B une deuxième matrice symplectique, alors :

$$(AB)^T \mathcal{J} AB = (B)^T (A)^T \mathcal{J} AB = (B)^T \mathcal{J} B = \mathcal{J},$$

donc le produit AB est une matrice symplectique et on a bien une structure de groupe multiplicatif.

Considérons finalement une matrice A symplectique, alors :

$$\begin{aligned} (A)^T \mathcal{J} A = \mathcal{J} &\Rightarrow \mathcal{J}(A)^T \mathcal{J} A = \mathcal{J}^2 = -\mathbb{I} \Rightarrow A \mathcal{J}(A)^T \mathcal{J} A = -A \\ &\Rightarrow A \mathcal{J}(A)^T \mathcal{J} A A^{-1} = -\mathbb{I} \Rightarrow A \mathcal{J}(A)^T \mathcal{J} = -\mathbb{I} \\ &\Rightarrow A \mathcal{J}(A)^T \mathcal{J} \mathcal{J}^{-1} = -\mathcal{J}^{-1} \Rightarrow A \mathcal{J}(A)^T = \mathcal{J}, \end{aligned}$$

c'est-à-dire que $(A)^T$ est une matrice symplectique. \square

EXEMPLE 2.6 (Le groupe $Sp(1, \mathbb{R})$). *Le groupe $Sp(1, \mathbb{R})$ des matrices symplectiques 2×2 coïncide avec le groupe $SL(2, \mathbb{R})$ des matrices 2×2 avec déterminant 1.*

En effet soit $A \in Sp(1, \mathbb{R})$, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ alors la condition (III.2.7) devient :

$$(A)^T \mathcal{J} A = \begin{pmatrix} 0 & -ad + bc \\ -bc + ad & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

c'est-à-dire $ad - bc = 1$.

EXEMPLE 2.7 (Le groupe $Sp(m, \mathbb{R})$, $m > 1$). *Soit A une matrice symplectique $2m \times 2m$ et écrivons-la comme matrice à blocs $m \times m$:*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad a, b, c \text{ et } d \in \mathbb{R}^{m \times m}.$$

La condition (III.2.7) devient alors une contrainte sur les matrices a, b, c et d :

$$(A)^T \mathcal{J} A = \begin{pmatrix} -(a)^T c + (c)^T a & -(a)^T d + (c)^T b \\ -(b)^T c + (d)^T a & -(b)^T d + (d)^T b \end{pmatrix} = \begin{pmatrix} \emptyset & -\mathbb{I} \\ \mathbb{I} & \emptyset \end{pmatrix},$$

donc A est symplectique si et seulement si :

$$(a)^T d - (c)^T b = \mathbb{I}, \quad (a)^T c \text{ et } (b)^T d \text{ sont des matrices symétriques.}$$

Nous avons donc en dimension $m > 1$ une condition plus restrictive que dans le cas $m = 1$.

Nous pouvons maintenant établir le lien entre les groupes $Sp(m, \mathbb{R})$ et $sp(m, \mathbb{R})$.

PROPOSITION 2.8. *L'espace tangent à $Sp(m, \mathbb{R})$ en \mathbb{I} est l'espace des matrices hamiltoniennes :*

$$(III.2.8) \quad T_{\mathbb{I}}Sp(l, \mathbb{R}) = sp(m, \mathbb{R}).$$

DÉMONSTRATION. Soit $A(t)$ une courbe dans $Sp(m, \mathbb{R})$ qui passe par le point \mathbb{I} en $t = 0$, c'est-à-dire une matrice telle que :

$$(III.2.9) \quad (A(t))^T \mathcal{J} A(t) = \mathcal{J},$$

pour tout t dans un voisinage de zéro, et $A(0) = \mathbb{I}$. Pour déterminer l'espace tangent à $Sp(m, \mathbb{R})$ en \mathbb{I} nous devons dériver la relation (III.2.9) et l'évaluer en $t = 0$:

$$(\dot{A}(t))^T \mathcal{J} A(t) + (A(t))^T \mathcal{J} \dot{A}(t) = 0,$$

car la matrice \mathcal{J} est une constante. Posons par simplicité $B = \dot{A}(0)$; par définition $B \in T_{\mathbb{I}}Sp(m, \mathbb{R})$, mais B vérifie aussi la relation suivante :

$$(B)^T \mathcal{J} \mathbb{I} + (\mathbb{I})^T \mathcal{J} B = 0 \Rightarrow (B)^T \mathcal{J} + \mathcal{J} B = 0,$$

c'est-à-dire B est hamiltonienne. \square

De plus contre il est facile de démontrer qu'à toute matrice hamiltonienne correspond une courbe dans $Sp(m, \mathbb{R})$:

PROPOSITION 2.9. *Soit B une matrice hamiltonienne, alors la matrice $A(t) = e^{Bt}$ est symplectique pour tout $t \in \mathbb{R}$.*

DÉMONSTRATION. Par définition

$$e^{Bt} = \sum_{l \geq 0} \frac{t^l B^l}{l!},$$

et donc on peut facilement montrer que :

$$(e^{Bt})^T = e^{(B)^T t} \quad \text{et} \quad (e^{Bt})^{-1} = e^{-Bt}.$$

En utilisant le fait que B est hamiltonienne, $(B)^T \mathcal{J} + \mathcal{J} B = 0$, calculons maintenant :

$$e^{(B)^T t} \mathcal{J} = \sum_{l \geq 0} \frac{t^l ((B)^T)^l}{l!} \mathcal{J} = \sum_{l \geq 0} \frac{t^l ((B)^T)^{(l-1)}}{l!} (-\mathcal{J} B) = \sum_{l \geq 0} \frac{t^l ((B)^T)^{(l-2)}}{l!} (-1)^2 \mathcal{J} B^2,$$

et donc en itérant l'idée nous allons obtenir :

$$e^{(B)^T t} \mathcal{J} = \mathcal{J} \sum_{l \geq 0} \frac{t^l (-B)^l}{l!} = \mathcal{J} e^{-Bt}.$$

Mais alors nous avons :

$$(e^{Bt})^T \mathcal{J} - \mathcal{J} e^{-Bt} = 0 \Rightarrow (e^{Bt})^T \mathcal{J} e^{Bt} = \mathcal{J},$$

qui est la condition pour avoir $A = e^{Bt}$ symplectique. \square

Observons que ce résultat n'est pas étonnant vu dans l'optique des algèbres de Lie : un groupe de Lie peut être construit en prenant l'exponentielle de son algèbre de Lie.

3. Champs Hamiltoniens

Nous avons vu au Chapitre précédent la formulation des équations de Hamilton en forme compacte :

$$\dot{\vec{x}} = \mathcal{J} \nabla \mathcal{H},$$

où $\vec{x} \in \mathbb{R}^{2m}$, nous nous demandons maintenant quelle est la forme la plus générale d'un champ de vecteurs hamiltonien, ou de façon équivalente, étant donné un champ de vecteurs $\vec{X}(\vec{x}, t)$, défini sur \mathbb{R}^{2m+1} , sous quelles conditions ce champ est-il hamiltonien.

Un champs hamiltonien est défini comme suit :

DÉFINITION 3.1. *Un champ vectoriel $\vec{X}(\vec{x}, t)$ (ou une équation différentielle $\dot{\vec{x}} = \vec{X}(\vec{x}, t)$), défini sur \mathbb{R}^{2m+1} est hamiltonien, s'il existe une fonction $f : \mathbb{R}^{2m+1} \rightarrow \mathbb{R}$, de classe \mathcal{C}^2 telle que :*

$$(III.3.1) \quad \vec{X}(\vec{x}, t) = \mathcal{J} \nabla f(\vec{x}, t).$$

Dans ce cas la fonction f est appelée hamiltonien correspondant au champ \vec{X} , et le champ est dit gradient symplectique de f .

Cette définition ne permet pas de savoir si cette fonction hamiltonienne existe ou non, ni de la construire explicitement ; pour cela le résultat suivant est donc très utile :

THÉORÈME 3.2. *Une condition nécessaire et suffisante pour qu'un champ de vecteurs $\vec{X}(\vec{x}, t)$, défini sur \mathbb{R}^{2m+1} , soit hamiltonien, est que la matrice Jacobienne $\nabla_{\vec{x}} \vec{X}$ soit une matrice hamiltonienne pour tout (t, \vec{x}) .*

DÉMONSTRATION. La condition est nécessaire, car si le champ est hamiltonien et si f est la fonction d'Hamilton associée alors nous pouvons calculer la Jacobienne du champ de vecteurs :

$$\frac{\partial X_i}{\partial x_j} = \frac{\partial}{\partial x_j} \sum_{l=1}^{2m} \mathcal{J}_{il} \frac{\partial f}{\partial x_l} = \sum_{l=1}^{2m} \mathcal{J}_{il} \frac{\partial^2 f}{\partial x_j \partial x_l} = (\mathcal{J} S)_{ij} \quad \forall i, j \in \{1, \dots, 2m\},$$

c'est-à-dire la matrice $A = \frac{\partial X}{\partial x}$, dont les composantes sont : $A_{ij} = \frac{\partial X_i}{\partial x_j}$, est de la forme :

$$A = \mathcal{J} S,$$

où la matrice $S = \frac{\partial^2 f}{\partial x \partial x}$ est symétrique. Donc par le point (2) du Théorème 2.3 la matrice A est hamiltonienne.

Montrons maintenant que la condition est aussi suffisante. Si $A = \frac{\partial X}{\partial x}$ est hamiltonienne pour tout (t, \vec{x}) , alors si on définit $\vec{Y} = \mathcal{J} \vec{X}$, la matrice $B = \frac{\partial Y}{\partial x}$ est symétrique (Théorème 2.3 point (3)), c'est-à-dire :

$$\frac{\partial Y_i}{\partial x_j} = \frac{\partial Y_j}{\partial x_i}.$$

Il existe donc une fonction f, \mathcal{C}^2 , telle que :

$$\vec{Y}(\vec{x}, t) = -\nabla f(\vec{x}, t),$$

et finalement

$$\vec{X} = \mathcal{J}^{-1} \vec{Y} = -\mathcal{J}(-\nabla f(\vec{x}, t)) = \mathcal{J} \nabla f(\vec{x}, t).$$

□

EXEMPLE 3.3. Considérons le système d'équations différentielles :

$$(III.3.2) \quad \begin{cases} \dot{p} &= -p^{\alpha+1}q^\delta \\ \dot{q} &= p^\alpha q^\beta, \end{cases}$$

et cherchons les conditions sur les paramètres α, β et γ pour qu'il soit hamiltonien.

Le champ de vecteurs associé est : $\vec{X}(\vec{x}) = (-x_1^{\alpha+1}x_2^\delta, x_1^\alpha x_2^\beta)$, où nous avons posé $x_1 = p$ et $x_2 = q$. Calculons la Jacobienne du champ de vecteurs :

$$\begin{aligned} \frac{\partial X_1}{\partial x_1} &= -(\alpha+1)x_1^\alpha x_2^\delta \\ \frac{\partial X_1}{\partial x_2} &= -(\delta)x_1^{\alpha+1}x_2^{\delta-1} \\ \frac{\partial X_2}{\partial x_1} &= \alpha x_1^{\alpha-1}x_2^\beta \\ \frac{\partial X_2}{\partial x_2} &= \beta x_1^\alpha x_2^{\beta-1}. \end{aligned}$$

La matrice $\frac{\partial X}{\partial x}$ est hamiltonienne si et seulement si pour tout x_1 et x_2 :

$$0 = \left(\frac{\partial X}{\partial x}\right)^T \mathcal{J} + \mathcal{J} \frac{\partial X}{\partial x} = \begin{pmatrix} 0 & (\alpha+1)x_1^\alpha x_2^\delta - \beta x_1^\alpha x_2^{\beta-1} \\ -(\alpha+1)x_1^\alpha x_2^\delta + \beta x_1^\alpha x_2^{\beta-1} & 0 \end{pmatrix}.$$

La seule possibilité est donc de vérifier les relations suivantes entre les paramètres :

$$\begin{aligned} \alpha &= -1, \quad \beta = 0 \quad \text{et } \delta \text{ quelconque} \\ \text{ou } \alpha + 1 &= \beta \quad \text{et } \delta = \beta - 1. \end{aligned}$$

Dans le cas générique le système devient :

$$\begin{cases} \dot{p} &= -p^{\alpha+1}q^\alpha \\ \dot{q} &= p^\alpha q^{\alpha+1}. \end{cases}$$

La méthode ne nous donne pas la fonction d'Hamilton associée au système, utilisons donc une deuxième méthode. Si le système (III.3.2) était hamiltonien il devrait exister une fonction $\mathcal{H}(p, q)$, telle que :

$$\begin{cases} \dot{p} &= -\frac{\partial \mathcal{H}}{\partial q} \\ \dot{q} &= \frac{\partial \mathcal{H}}{\partial p}, \end{cases}$$

alors en imposant la seconde équation nous devons avoir

$$\frac{\partial \mathcal{H}}{\partial p} = p^\alpha q^\beta,$$

qui peut être intégrée par rapport à p :

$$\mathcal{H} = \frac{1}{\alpha+1} p^{\alpha+1} q^\beta + \phi(q) \quad \text{si } \alpha \neq -1,$$

avec $\phi(q)$ fonction arbitraire de la variable q , ou bien :

$$\mathcal{H} = (\log p) q^\beta + \phi(q) \quad \text{si } \alpha = -1.$$

Dans les deux cas la fonction ϕ est déterminée en imposant que la dérivée par rapport à q de la fonction \mathcal{H} coïncide avec la première équation (III.3.2). Donc par

un calcul élémentaire :

$$\begin{aligned} -p^{\alpha+1}q^\delta &= -\frac{\beta}{\alpha+1}p^{\alpha+1}q^{\beta-1} - \phi'(q) \text{ si } \delta \neq -1 \\ -p^{\alpha+1}q^\delta &= -(\log p)\beta q^{\beta-1} - \phi'(q) \text{ si } \delta = -1. \end{aligned}$$

Dans le deuxième cas la relation peut être vérifiée si et seulement si :

$$\beta = 0, \quad \alpha + 1 = 0, \quad \delta \text{ quelconque et } \phi'(q) = -q^\delta,$$

par contre dans le cas $\delta \neq -1$ la relation est toujours satisfaite si et seulement si :

$$\frac{\beta}{\alpha+1} = 1 \quad \text{et} \quad \delta = \beta - 1,$$

qui sont les mêmes conditions que celles déterminées précédemment.

4. Transformations canoniques

Une façon possible de résoudre une équation différentielle est de déterminer des coordonnées bien choisies où l'équation différentielle assume une forme "plus simple". En général si nous avons l'EDO :

$$(III.4.1) \quad \dot{\vec{x}} = \vec{v}(\vec{x}, t),$$

alors le changement de coordonnées $\vec{x} \mapsto \vec{y}$, régulier et inversible, donné par la transformation :

$$(III.4.2) \quad \vec{x} = \vec{x}(\vec{y}, t),$$

permet de récrire l'EDO (III.4.1) dans la forme suivante ¹ :

$$(III.4.3) \quad \dot{\vec{y}} = \vec{w}(\vec{y}, t),$$

où le champ de vecteurs \vec{w} est donné par :

$$\vec{w}(\vec{y}, t) = A\vec{v}(\vec{x}(\vec{y}, t), t) + \frac{\partial \vec{y}}{\partial t},$$

avec A nous avons dénoté la Jacobienne de la transformation inverse, c'est-à-dire :

$$A_{ij} = \frac{\partial y_i}{\partial x_j}.$$

Si nous supposons que le champ du départ est hamiltonien :

$$(III.4.4) \quad \dot{\vec{x}} = \mathcal{J}\nabla\mathcal{H}(\vec{x}, t),$$

pour une certaine fonction régulière $\mathcal{H} : \mathbb{R}^{2m+1} \rightarrow \mathbb{R}$, alors généralement dans les nouvelles coordonnées \vec{y} , l'EDO n'est pas hamiltonienne, c'est-à-dire ce n'est pas dit qu'il existe une fonction $\mathcal{K}(\vec{y}, t)$ telle que :

$$(III.4.5) \quad \mathcal{J}\nabla_{\vec{y}}\mathcal{K}(\vec{y}, t) = A\mathcal{J}\nabla_{\vec{x}}\mathcal{H}(\vec{x}, t) + \frac{\partial \vec{y}}{\partial t}.$$

¹On dérive la relation inverse de (III.4.2) par rapport au temps et on obtient :

$$\dot{y}_i = \frac{d}{dt}y_i(\vec{x}, t) = \sum_{j=1}^n \frac{\partial y_i}{\partial x_j} \frac{dx_j}{dt} + \frac{\partial y_i}{\partial t} = \sum_{j=1}^n A_{ij}v_j + \frac{\partial y_i}{\partial t}.$$

EXEMPLE 4.1. Reprenons l'exemple 2.1 d'un système hamiltonien linéaire $\mathcal{H} = \frac{1}{2}(\vec{x})^T S \vec{x}$, avec S une quelconque matrice symétrique, et effectuons une transformation linéaire :

$$\vec{y} = A \vec{x},$$

avec A matrice $2m \times 2m$, inversible. Alors les équations d'Hamilton (cfr. exemple 2.1) :

$$\dot{\vec{x}} = \mathcal{J} S \vec{x},$$

se transforment en :

$$\dot{\vec{y}} = A \mathcal{J} S A^{-1} \vec{y}.$$

Or ces équations sont hamiltoniennes si et seulement si il existe une matrice symétrique B telle que :

$$(III.4.6) \quad \mathcal{J} B = A \mathcal{J} S A^{-1},$$

quelle que soit la matrice S symétrique.

La relation (III.4.6) peut être réécrite sous la forme équivalente :

$$(A)^T B A = -(A)^T \mathcal{J} A \mathcal{J} S = -\Lambda \mathcal{J} S,$$

avec $\Lambda = (A)^T \mathcal{J} A$, telle que :

$$(\Lambda)^T = -\Lambda.$$

L'existence de B symétrique est donc équivalent à la condition de symétrie suivante :

$$(III.4.7) \quad ((A)^T B A)^T = (A)^T (B)^T A = (A)^T B A \Rightarrow -S \mathcal{J} \Lambda = -\Lambda \mathcal{J} S,$$

Si A est symplectique alors (en réalité ça suffit $\Lambda = a \mathcal{J}$ avec $a \neq 0$) :

$$\Lambda = (A)^T \mathcal{J} A = \mathcal{J},$$

et la relation (III.4.7) est vérifiée quelle que soit la matrice S . En réalité cette condition est aussi nécessaire. Si nous écrivons les matrices S et Λ en blocs $m \times m$:

$$S = \begin{pmatrix} \alpha & \beta \\ (\beta)^T & \gamma \end{pmatrix} \quad \text{et} \quad \Lambda = \begin{pmatrix} \lambda & \mu \\ -(\mu)^T & \nu \end{pmatrix},$$

avec les conditions $(\alpha)^T = \alpha$, $(\gamma)^T = \gamma$, $(\lambda)^T = -\lambda$ et $(\nu)^T = -\nu$. Alors la condition (III.4.7) devient :

$$\begin{cases} -\lambda(\beta)^T + \mu\alpha &= \alpha(\mu)^T + \beta\lambda \\ -\lambda\gamma + \mu\beta &= -\alpha\nu + \beta\mu \\ (\mu)^T(\beta)^T + \nu\alpha &= (\beta)^T(\mu)^T + \gamma\lambda \\ (\mu)^T\gamma + \nu\beta &= -(\beta)^T\nu + \gamma\mu; \end{cases}$$

or car S est arbitraire nous allons choisir différents cas "spéciaux" :

- si $\alpha = \gamma = 0$, alors μ doit commuter (deuxième équation) avec toute matrice $m \times m$, β , et donc $\mu = a \mathbb{I}$, avec $a \neq 0$;
- si $\alpha = \beta = 0$, alors $\lambda = 0$ (troisième équation);
- si $\beta = \gamma = 0$, alors $\nu = 0$ (troisième équation).

Mais alors

$$\Lambda = a \begin{pmatrix} \emptyset & \mathbb{I} \\ -\mathbb{I} & \emptyset \end{pmatrix} = a \mathcal{J},$$

de plus si $a = 1$ le nouvel hamiltonien est simplement le transformé du vieux :

$$\mathcal{K}(\vec{y}) = \mathcal{H}(\vec{x}).$$

Il est évident que la conservation de la structure des équations d'Hamilton est très importante, surtout dans les applications comme la *Théorie des perturbations*, nous allons donc étudier maintenant les hypothèses sous lesquelles la forme des équations d'Hamilton est préservée par une transformation de coordonnées. Nous commençons par formaliser ce qu'on entend par "conservation" des équations d'Hamilton.

DÉFINITION 4.2. Une transformation de coordonnées $\vec{x} = \vec{x}(\vec{y}, t)$, régulière et inversible pour tout t fixé, préserve la structure canonique des équations d'Hamilton, si pour tout hamiltonien $\mathcal{H}(\vec{x}, t)$, il existe une fonction, $\mathcal{K}(\vec{y}, t)$, dite nouvel hamiltonien, telle que dans ces nouvelles coordonnées les équations du mouvement sont dans la forme hamiltonienne, $\vec{y} = (\vec{P}, \vec{Q})$:

$$\begin{cases} \dot{P}_j &= -\frac{\partial \mathcal{K}}{\partial Q_j} \\ \dot{Q}_j &= \frac{\partial \mathcal{K}}{\partial P_j} \quad \forall j = 1, \dots, m. \end{cases}$$

EXEMPLE 4.3. Soit $a(t)$ une fonction différentiable positive, alors la transformation :

$$\vec{Q} = a(t)\vec{q} \quad \text{et} \quad \vec{P} = \frac{1}{a(t)}\vec{p},$$

préserve la structure canonique.

Par dérivation du changement de coordonnées par rapport au temps, on trouve les équations du mouvement dans les nouvelles variables :

$$(III.4.8) \quad \begin{cases} \dot{\vec{Q}} &= \dot{a}(t)\vec{q} + a(t)\frac{\partial \mathcal{H}}{\partial \vec{p}} = \frac{\dot{a}(t)}{a(t)}\vec{Q} + a(t)\frac{\partial \mathcal{H}}{\partial \vec{p}} \\ \dot{\vec{P}} &= -\frac{1}{a^2(t)}\dot{a}(t)\vec{p} - \frac{1}{a(t)}\frac{\partial \mathcal{H}}{\partial \vec{q}} = -\frac{1}{a(t)}\dot{a}(t)\vec{P} - \frac{1}{a(t)}\frac{\partial \mathcal{H}}{\partial \vec{q}}, \end{cases}$$

mais on peut facilement vérifier que ces équations correspondent aux équations d'Hamilton pour le nouvel Hamiltonien \mathcal{K} :

$$\mathcal{K}(\vec{P}, \vec{Q}, t) := \mathcal{H}\left(a(t)\vec{P}, \frac{\vec{Q}}{a(t)}, t\right) + \frac{\dot{a}(t)}{a(t)}\vec{P} \cdot \vec{Q}.$$

En effet :

$$\begin{cases} \dot{\vec{P}} &= -\frac{\partial \mathcal{K}}{\partial \vec{Q}} = -\frac{\partial \mathcal{H}}{\partial \vec{q}} \frac{\partial \vec{q}}{\partial \vec{Q}} - \frac{\dot{a}(t)}{a(t)}\vec{P} = -\frac{1}{a(t)}\frac{\partial \mathcal{H}}{\partial \vec{q}} - \frac{\dot{a}(t)}{a(t)}\vec{P} \\ \dot{\vec{Q}} &= \frac{\partial \mathcal{K}}{\partial \vec{P}} = \frac{\partial \mathcal{H}}{\partial \vec{p}} \frac{\partial \vec{p}}{\partial \vec{P}} + \frac{\dot{a}(t)}{a(t)}\vec{Q} = a(t)\frac{\partial \mathcal{H}}{\partial \vec{p}} + \frac{\dot{a}(t)}{a(t)}\vec{Q}, \end{cases}$$

qui coïncident avec (III.4.8).

EXEMPLE 4.4. La transformation qui change, à un signe près, les coordonnées \vec{q} avec les moments cinétiques \vec{p} , $\vec{P} = -\vec{q}$ et $\vec{Q} = \vec{p}$, préserve la structure canonique ; le nouvel hamiltonien est lié au vieux par :

$$\mathcal{K}(\vec{P}, \vec{Q}, t) = \mathcal{H}(\vec{Q}, -\vec{P}, t).$$

Dans la mécanique lagrangienne, les coordonnées \vec{q} , étaient grosso modo liées à des positions (rectangulaire, polaires ou généralisées) et les \vec{p} étaient des vitesses. Cet exemple montre comme dans le formalisme hamiltonien le rôle de ces coordonnées peut être interchangé.

EXEMPLE 4.5. Supposons qu'on ait un changement de coordonnées qui agit sur les \vec{q} seulement : $\vec{Q} = \vec{Q}(\vec{q})$. Nous cherchons comment les variables \vec{p} doivent changer pour préserver la structure des équations d'Hamilton avec cette transformation.

Les vitesses généralisées sont obtenues par dérivation par rapport à t du changement de coordonnées :

$$\dot{Q}_j = \sum_{l=1}^m \frac{\partial Q_j}{\partial q_l} \dot{q}_l \Rightarrow \dot{\vec{Q}} = A(\vec{q}) \dot{\vec{q}},$$

où nous avons dénoté par $A(\vec{q})$ la matrice Jacobienne du changement de coordonnées. Si $\mathcal{L}(\vec{q}, \dot{\vec{q}}, t)$ est la fonction de Lagrange du système, alors nous pouvons introduire le lagrangien dans les nouvelles coordonnées, $\hat{\mathcal{L}}$:

$$\hat{\mathcal{L}}(\vec{Q}, \dot{\vec{Q}}, t) := \mathcal{L}(\vec{q}(\vec{Q}), A^{-1}(\vec{q}(\vec{Q})) \dot{\vec{Q}}, t).$$

Par définition le moment cinétique est obtenu en dérivant le lagrangien par rapport aux vitesses généralisées, donc :

$$P_j = \frac{\partial \hat{\mathcal{L}}}{\partial \dot{Q}_j} = \sum_{l=1}^m \frac{\partial \mathcal{L}}{\partial \dot{q}_l} \frac{\partial \dot{q}_l}{\partial \dot{Q}_j} = \sum_{l=1}^m p_l (A^{-1})_{lj}.$$

La transformation sur les variables \vec{q} "induit" une transformation sur les variables \vec{p} :

$$\vec{P} = ((A)^T)^{-1} \vec{p}.$$

Pour un hamiltonien donné, les équations du mouvement dans les nouvelles coordonnées sont ² :

$$\begin{cases} \dot{P}_j &= \dot{p}_l (A^{-1})_{lj} + p_l \left(\frac{\partial (A^{-1})_{lj}}{\partial q_k} \dot{q}_k \right) = -\frac{\partial \mathcal{H}}{\partial q_l} (A^{-1})_{lj} + A_{hl} P_h \frac{\partial (A^{-1})_{lj}}{\partial q_k} (A^{-1})_{kn} \dot{Q}_n \\ &= -\frac{\partial \mathcal{H}}{\partial q_l} (A^{-1})_{lj} + A_{hl} P_h \frac{\partial (A^{-1})_{lj}}{\partial q_k} (A^{-1})_{kn} A_{ns} \frac{\partial \mathcal{H}}{\partial p_s} = -\frac{\partial \mathcal{H}}{\partial q_l} (A^{-1})_{lj} + A_{hl} P_h \frac{\partial (A^{-1})_{lj}}{\partial q_k} \frac{\partial \mathcal{H}}{\partial p_k} \\ \dot{Q}_j &= A_{jl} \dot{q}_l = A_{jl} \frac{\partial \mathcal{H}}{\partial p_l}. \end{cases}$$

Nous nous demandons si ces équations proviennent d'un nouvel hamiltonien ; pour cela nous allons calculer l'hamiltonien associé par transformation de Legendre du nouveau lagrangien :

$$\hat{\mathcal{H}}(\vec{P}, \vec{Q}, t) = \vec{P} \cdot \dot{\vec{Q}} - \hat{\mathcal{L}}(\vec{Q}, \dot{\vec{Q}}, t),$$

or nous observons que (encore convention sommation indices répétés) :

$$\vec{P} \cdot \dot{\vec{Q}} = ((A)^T)^{-1} \vec{p} \cdot A \dot{\vec{q}} = (A^{-1})_{jl} p_j A_{lk} \dot{q}_k = p_j \dot{q}_j = \vec{p} \cdot \dot{\vec{q}},$$

où nous avons utilisé $(A^{-1})_{jl} A_{lk} = \delta_{jk}$. Donc le nouveau hamiltonien est le transformé du vieux :

$$\hat{\mathcal{H}}(\vec{P}, \vec{Q}, t) = \mathcal{H}(\vec{p}, \vec{q}, t) = \mathcal{H}((A)^T \vec{P}, \vec{q}(\vec{Q}), t),$$

et donc la structure canonique est préservée.

Nous pouvons maintenant définir

DÉFINITION 4.6 (Transformation canonique). Une transformation de coordonnées $\vec{y} = \vec{y}(\vec{x}, t)$, différentiable et inversible pour tout t fixé est dite canonique si la matrice Jacobienne :

$$A(\vec{x}, t)_{ij} = \frac{\partial y_i}{\partial x_j},$$

²Pour abréger les notations, nous utilisons ici la convention de sommation sur les indices répétés, par exemple $a_i b_i \equiv \sum_{i=1}^m a_i b_i = \vec{a} \cdot \vec{b}$ ou $A_{ij} a_j \equiv \sum_{j=1}^m A_{ij} a_j = A \vec{a}$.

est symplectique pour tout (\vec{x}, t) dans le domaine de définition de la transformation. Si la transformation canonique est indépendante du temps, $\vec{y} = \vec{y}(\vec{x})$, elle sera appelée complètement canonique.

Observons que la Jacobienne d'une composition de transformations est le produit des Jacobiennes, mais alors puisque les matrices symplectiques forment un groupe par rapport au produit de matrices, nous pouvons conclure que les transformations canoniques forment un groupe par composition. Le résultat suivant nous permet de répondre à la question soulevée au début de cette section, à savoir : quelles transformations de coordonnées préservent les équations d'Hamilton.

THÉOREME 4.7. *Les transformations canoniques préservent la structure des équations d'Hamilton.*

DÉMONSTRATION. Soit donc la transformation $\vec{y} = \vec{y}(\vec{x}, t)$ et dérivons-la par rapport au temps :

$$(III.4.9) \quad \dot{\vec{y}} = \frac{\partial \vec{y}}{\partial t} + \frac{\partial \vec{y}}{\partial \vec{x}} \dot{\vec{x}}.$$

Dans les variables du départ les équations du mouvement sont dans la forme hamiltonienne canonique, $\dot{\vec{x}} = \mathcal{J} \nabla_{\vec{x}} \mathcal{H}(\vec{x}, t)$, et donc dans les nouvelles variables nous avons :

$$\dot{\vec{y}} = \frac{\partial \vec{y}}{\partial t} + A \mathcal{J} \nabla_{\vec{x}} \mathcal{H}(\vec{x}, t),$$

où nous avons dénoté par A la matrice Jacobienne.

Définissons le transformé du vieux hamiltonien :

$$\hat{\mathcal{H}}(\vec{y}, t) := \mathcal{H}(\vec{x}(\vec{y}, t), t),$$

où la transformation $\vec{x} = \vec{x}(\vec{y}, t)$ est l'inverse du changement de variables du départ. Alors nous pouvons calculer :

$$\frac{\partial \hat{\mathcal{H}}}{\partial y_j} = \frac{\partial \mathcal{H}}{\partial x_i} \frac{\partial x_k}{\partial y_j} = \frac{\partial \mathcal{H}}{\partial x_k} (A^{-1})_{kj},$$

en multipliant par A_{jh} nous obtenons :

$$A_{jh} \frac{\partial \hat{\mathcal{H}}}{\partial y_j} = \frac{\partial \mathcal{H}}{\partial x_k} (A^{-1})_{kj} A_{jh} = \frac{\partial \mathcal{H}}{\partial x_k} \Rightarrow (A)^T \nabla_{\vec{y}} \hat{\mathcal{H}} = \nabla_{\vec{x}} \mathcal{H}.$$

Ce qui nous permet de récrire (III.4.9) de la façon suivante :

$$(III.4.10) \quad \dot{\vec{y}} = \frac{\partial \vec{y}}{\partial t} + A \mathcal{J} (A)^T \nabla_{\vec{y}} \hat{\mathcal{H}}(\vec{y}, t),$$

et grâce à l'hypothèse de transformation canonique, nous avons³ : $A \mathcal{J} = \mathcal{J} ((A)^T)^{-1}$ et donc :

$$(III.4.11) \quad \dot{\vec{y}} = \frac{\partial \vec{y}}{\partial t} + \mathcal{J} ((A)^T)^{-1} (A)^T \nabla_{\vec{y}} \hat{\mathcal{H}}(\vec{y}, t) = \frac{\partial \vec{y}}{\partial t} + \mathcal{J} \nabla_{\vec{y}} \hat{\mathcal{H}}(\vec{y}, t).$$

Pour conclure la démonstration il nous reste à prouver que le champ $\frac{\partial \vec{y}}{\partial t}$ est lui aussi hamiltonien. Grâce au Théorème 3.2 cela est équivalent à montrer que la matrice

³Un calcul élémentaire nous donne :

$$(A)^T \mathcal{J} A = \mathcal{J} \Rightarrow (A)^T \mathcal{J} A \mathcal{J} = -\mathbb{I} \Rightarrow A \mathcal{J} = \mathcal{J} ((A)^T)^{-1}.$$

Jacobienne de ce champ, $B = \nabla_{\vec{y}} \frac{\partial \vec{y}(\vec{y}, t, t)}{\partial t}$, est hamiltonienne. Un simple calcul montre que :

$$(III.4.12) \quad B = \frac{\partial A}{\partial t} A^{-1},$$

en fait :

$$B_{ij} = \frac{\partial}{\partial y_j} \frac{\partial y_i(\vec{y}, t, t)}{\partial t} = \frac{\partial^2 y_i(\vec{y}, t, t)}{\partial t \partial x_k} \frac{\partial x_k}{\partial y_j} = \frac{\partial}{\partial t} \frac{\partial y_i(\vec{y}, t, t)}{\partial x_k} \frac{\partial x_k}{\partial y_j} = \frac{\partial A_{ik}}{\partial t} A_{kj}^{-1},$$

car B dépend du temps de façon explicite mais aussi *via* la fonction $\vec{y}(\vec{y}, t)$. Il nous reste à prouver que si A est symplectique alors B donné par (III.4.12) est hamiltonienne, et pour le Théorème 2.3 il suffit de prouver que la matrice : $C = \mathcal{J}B$ est symétrique.

Si on dérive la relation $(A)^T \mathcal{J}A = \mathcal{J}$ par rapport à t nous obtenons :

$$\frac{\partial(A)^T}{\partial t} \mathcal{J}A + (A)^T \mathcal{J} \frac{\partial A}{\partial t} = 0,$$

en multipliant à gauche par $(A^{-1})^T$ et à droite par A^{-1} nous trouvons :

$$(A^{-1})^T \frac{\partial(A)^T}{\partial t} \mathcal{J}A A^{-1} + (A^{-1})^T (A)^T \mathcal{J} \frac{\partial A}{\partial t} A^{-1} = 0 \Rightarrow (B)^T \mathcal{J} + \mathcal{J}B = 0,$$

et finalement :

$$-(\mathcal{J}B)^T + \mathcal{J}B = 0 \Rightarrow (C)^T = C,$$

ce qui termine la démonstration. \square

REMARQUE 4.8. *Le Théorème nous donne aussi le nouvel hamiltonien, car d'après la démonstration nous avons :*

$$\mathcal{K} = \hat{\mathcal{H}} + \mathcal{K}_0,$$

où $\hat{\mathcal{H}}$ est le vieux hamiltonien récrit dans les nouvelles variables, et \mathcal{K}_0 est l'hamiltonien associé au champs de vecteurs hamiltonien $\frac{\partial \vec{y}}{\partial t}$, c'est-à-dire :

$$\frac{\partial \vec{y}}{\partial t} = \mathcal{J} \nabla_{\vec{y}} \mathcal{K}_0.$$

Si la transformation est complètement canonique, alors $\mathcal{K}_0 \equiv 0$ et le nouveau hamiltonien est le transformé du vieux.

Nous avons déjà vu une propriété importante du flot hamiltonien dans la section II.4, à savoir l'invariance des volumes dans l'espace des phases; ici nous présenterons une deuxième propriété remarquable : le *flot hamiltonien définit une transformation canonique*. Autrement dit la solution des équations d'Hamilton pour un hamiltonien donné, vu comme application à temps fixé, préserve la structure des équations d'Hamilton pour tout autre hamiltonien. Dans la section suivante nous interpréterons ce résultat dans son cadre naturel de la géométrie symplectique.

THÉOREME 4.9. *Soit \mathcal{H} un hamiltonien et soit Φ^t son flot pour t réel, c'est-à-dire pour tout t :*

$$(III.4.13) \quad \Phi^t(\vec{p}(0), \vec{q}(0)) = (\vec{p}(t), \vec{q}(t)),$$

où $(\vec{p}(t), \vec{q}(t))$ est la solution des équations d'Hamilton pour l'hamiltonien \mathcal{H} avec données initiales $(\vec{p}(0), \vec{q}(0))$.

Alors l'application (III.4.13) est canonique.

DÉMONSTRATION. Notons pour simplifier $\vec{x} = (\vec{p}(0), \vec{q}(0))$ et soit :

$$A_t(\vec{x}) := \frac{\partial \Phi^t}{\partial \vec{x}},$$

la matrice Jacobienne du flot pour un t fixé. Pour démontrer le résultat nous prouverons que A_t est symplectique pour tout (t, \vec{x}) , c'est-à-dire :

$$(A_t(\vec{x}))^T \mathcal{J} A_t(\vec{x}) = \mathcal{J}.$$

Nous remarquons qu'il est équivalent de démontrer que pour tout t on a :

$$(III.4.14) \quad \frac{\partial}{\partial t} ((A_t(\vec{x}))^T \mathcal{J} A_t(\vec{x})) = 0,$$

en effet cela implique que la matrice $(A_t(\vec{x}))^T \mathcal{J} A_t(\vec{x})$ est constante par rapport à t et en particulier pour $t = 0$ elle vaut \mathcal{J} , car $A_{t=0}(\vec{x}) = \mathbb{I}$ (si $t = 0$ aucun point n'est bougé par le flot).

Au lieu de démontrer (III.4.14) pour tout t nous montrons qu'il est suffisant de le prouver pour $t = 0$; le flot définit un groupe à un paramètre, c'est-à-dire :

$$\Phi^{(t+\tau)}(\vec{x}) = \Phi^\tau(\Phi^t(\vec{x})),$$

mais alors la matrice Jacobienne vérifie :

$$A_{t+\tau}(\vec{x}) = A_\tau(\Phi^t(\vec{x})) A_t(\vec{x}),$$

et donc (III.4.14) se réécrit comme (car si $\tau = 0$ alors $t + \tau = t$ arbitraire) :

$$\frac{\partial}{\partial \tau} ((A_{t+\tau}(\vec{x}))^T \mathcal{J} A_{t+\tau}(\vec{x})) = (A_t(\vec{x}))^T \frac{\partial}{\partial \tau} [(A_\tau(\Phi^t(\vec{x})))^T \mathcal{J} A_\tau(\Phi^t(\vec{x}))]_{\tau=0} A_t(\vec{x}) = 0.$$

Intéressons-nous donc à $t = 0$. Or

$$(III.4.15) \quad \left. \frac{\partial (A_t)_{ij}}{\partial t} \right|_{t=0} = \left. \frac{\partial}{\partial x_j} \frac{d\Phi_i^t}{dt} \right|_{t=0} = \frac{\partial}{\partial x_j} \mathcal{J}_{il} \frac{\partial \mathcal{H}}{\partial x_l}(\vec{x}) = \mathcal{J}_{il} \frac{\partial \mathcal{H}}{\partial x_l x_j}(\vec{x}),$$

(notation indices répétés sous-entendent sommation appliquée à $\vec{X} = \mathcal{J} \nabla \mathcal{H}$), donc en notation matricielle :

$$\left. \frac{\partial A_t}{\partial t} \right|_{t=0} = \mathcal{J} \text{Hes}(\mathcal{H}),$$

où $\text{Hes}(\mathcal{H})$ est la matrice Hessienne associée à l'hamiltonien \mathcal{H} , et de façon similaire :

$$\left. \frac{\partial (A_t)^T}{\partial t} \right|_{t=0} = \text{Hes}(\mathcal{H})(\mathcal{J})^T,$$

car la matrice Hessienne est symétrique. Nous pouvons donc conclure que :

$$\left. \frac{\partial}{\partial t} ((A_t)^T \mathcal{J} A_t) \right|_{t=0} = \text{Hes}(\mathcal{H})(\mathcal{J})^T \mathcal{J} A_{t=0} + (A_{t=0})^T \mathcal{J} \mathcal{J} \text{Hes}(\mathcal{H}) = \text{Hes}(\mathcal{H}) \mathbb{I} - \mathbb{I} \text{Hes}(\mathcal{H}) = 0,$$

où nous avons utilisé les relations : $(\mathcal{J})^T = \mathcal{J}^{-1}$, $\mathcal{J}^2 = -\mathbb{I}$ et $A_{t=0} = \mathbb{I}$ car à $t = 0$ le flot n'a déplacé aucun point et donc sa matrice Jacobienne est l'identité. \square

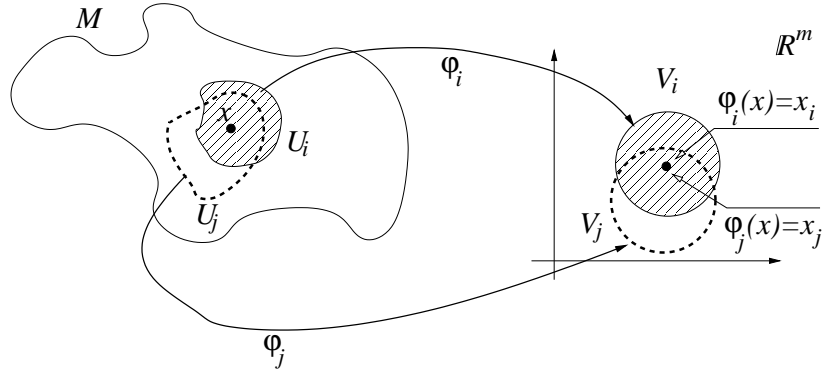


FIG. 2. Variété avec deux cartes locales, U_i et U_j , et les changements de coordonnées respectifs, φ_i et φ_j .

5. Fonctions, champs de vecteurs et formes différentielles sur variétés

Le but de cette section est de fournir un bref rappel des notions vues dans le cours de Géométrie différentielle, et donc supposées connues, liées aux variétés.

Soit M un variété différentiable de dimension m , alors il existe un système d'ouverts de la variété $U \subset M$ et des fonctions $\varphi : U \rightarrow V \subset \mathbb{R}^m$, régulières, inversible et avec inverse de la même régularité, c'est-à-dire un *atlas* (voir Fig. 2).

Si le point $x \in U_i \subset M$ on appellera son image par φ_i i -ème coordonnée : $\varphi_i(x) = x_i \in V_i$. Si deux cartes locales ont une intersection non vide, $U_i \cap U_j \neq \emptyset$, alors l'application (dite *fonction de transition*) $\psi_{ij} := \varphi_i \circ \varphi_j^{-1} : V_j \rightarrow V_i$ est un changement (local) de coordonnées dans \mathbb{R}^m . La régularité des fonctions ψ_{ij} détermine la régularité de la variété.

5.1. Fonctions. Une *fonction*, f , (ou *0-forme*) sur une variété M est un objet intrinsèque : $f : M \rightarrow \mathbb{R}$, tel que pour toute carte locale (U_i, φ_i) il existe une *représentation locale*, f_i :

$$f_i := f \circ \varphi_i^{-1} : V_i \rightarrow \mathbb{R},$$

c'est-à-dire une fonction "normale-usuelle". De plus pour tout couple de cartes locales, (U_i, φ_i) et (U_j, φ_j) , avec intersection non vide, $U_i \cap U_j \neq \emptyset$, alors on peut définir deux représentations locales, f_i et f_j , qui sont (c'est bien ici la contrainte que la fonction f soit un objet intrinsèque, c'est-à-dire défini indépendamment des coordonnées choisies) reliées par la relation (voir Fig. 3) :

$$(III.5.1) \quad f_j = f_i \circ \psi_{ij}.$$

5.2. Champs de vecteurs. Sur chaque point $x \in M$ il est bien défini un *espace tangent*, dénoté par $T_x M$, de dimension m . La réunion des tous les espaces tangents a une structure "naturelle" de variété appelée *fibré tangent*, $TM = \cup_{x \in M} T_x M$, dont la dimension est $2m$: un point de TM est un couple (x, \vec{X}) , avec $x \in M$ et $\vec{X} \in T_x M$. Un *champ de vecteurs* sur M est une application qui à tout point $x \in M$ associe un point de TM , $(x, \vec{X}(x))$, où le vecteur $\vec{X}(x)$ est dans $T_x M$.

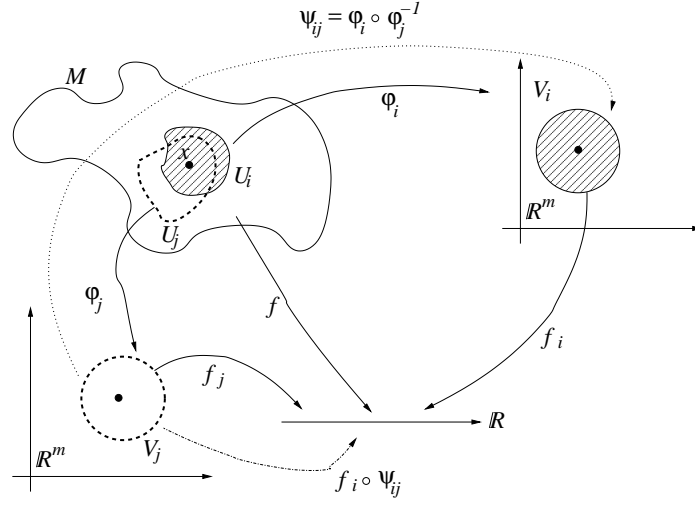


FIG. 3. Fonctions sur variété avec deux représentations locales.

Pour toute carte locale (U_i, φ_i) et pour tout point $x \in U_i \subset M$, nous pouvons définir une base de $T_x M$, formée par les vecteurs tangents aux lignes coordonnées (c'est-à-dire les images inverses par les fonctions φ_i des axes des coordonnées dans \mathbb{R}^m), notés généralement par $\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_m}$, parfois notés aussi pour abrégier $\partial_1, \dots, \partial_m$ (voir Fig. 4). Dans cette base le champ admet l'écriture suivante :

$$\vec{X}(x) = \sum_{i=1}^m X_i(x) \frac{\partial}{\partial x_i}.$$

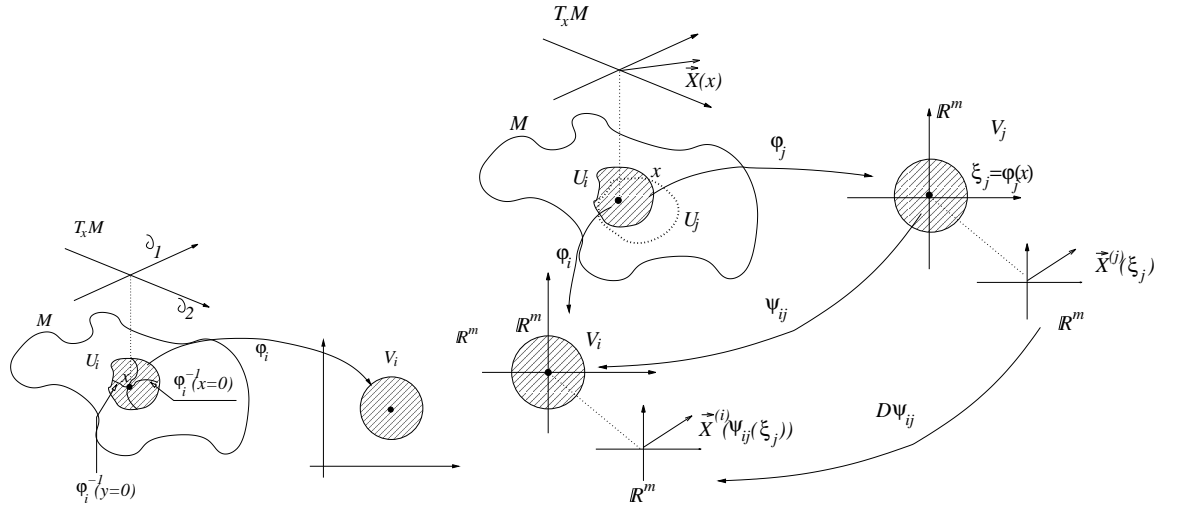


FIG. 4. Champs de vecteurs sur variété. À gauche construction de fibré tangent, à droite l'application tangente et la transformation des champs représentatifs locaux.

Soit (U, φ) une carte locale, alors pour tout $x \in U \subset M$, on définit l'*application tangente* (à φ) $D\varphi : T_x M \rightarrow \mathbb{R}^m$, où $\mathbb{R}^m \sim T_{\varphi(x)} V$. Par construction cette application associe à la base $\partial_1, \dots, \partial_m$ la base standard de \mathbb{R}^m , e_1, \dots, e_m .

Pour tout champ \vec{X} sur M , et pour toute carte (U_i, φ_i) , l'application tangente définit un champ local sur $V_i = \varphi_i(U_i)$ appelé *champ représentatif local* de \vec{X} :

$$\vec{X}(x) = \sum_{i=1}^m X_i(x) \frac{\partial}{\partial x_i} \Rightarrow_{D\varphi_i} \vec{X}^{(i)}(x) = \sum_{j=1}^m X_j(x) e_j,$$

où dans la formule de droite nous avons introduit un petit abus de langage en définissant $X_j(x)$, en réalité on devrait écrire $X_j(\varphi_i(x))$. Si $\vec{X}^{(i)}(x)$ et $\vec{X}^{(j)}(x)$ sont deux représentations locales du même champ avec deux cartes différentes (U_i, φ_i) et (U_j, φ_j) , avec $U_i \cap U_j \neq \emptyset$, alors ces deux champs locaux sont liés par la loi de transformation (qui rend l'objet champ intrinsèque) :

$$(III.5.2) \quad \vec{X}^{(j)}(\xi_j) = A^{-1} \vec{X}^{(i)}(\psi_{ij}(\xi_j)),$$

où $A = \frac{\partial \psi_{ij}}{\partial \xi_j}$ est la matrice Jacobienne de la fonction de transition et $\xi_j = \varphi_j(x)$ est l'image du point $x \in U_j \subset M$ dans V_j (i.e. coordonnées locales du point sur la variété).

Vice versa si nous avons une famille de champs locaux définis sur \mathbb{R}^m , $(\vec{X}^{(j)})_{1 \leq j \leq k}$, pour un certain k fini (plus grands que le nombre de cartes nécessaires pour paramétriser la variété M) qui vérifient les relations (III.5.2), alors il existe un champ \vec{X} global défini sur M , dont les $\vec{X}^{(j)}$ sont les champs représentatifs locaux.

5.3. 1-formes différentielles. Soit $x \in M$, comme nous avons pu définir l'espace tangent à M en x , nous pouvons aussi définir son *dual*, l'*espace cotangent*, noté par $T_x^* M$, c'est-à-dire l'ensemble des applications linéaires de $T_x M \rightarrow \mathbb{R}$. On peut montrer (voir cours de géométrie) que la réunion de tous les espaces cotangents a une structure de variété, appelée *fibré cotangent*, $T^* M = \cup_{x \in M} T_x^* M$, dont la dimension est $2m$ ⁴. Un point de $T^* M$ est un couple (x, θ) , où θ dépend de x mais nous ne l'avons pas noté pour simplifier les notations, avec $x \in M$ et $\theta \in T_x^* M$. Par définition de 1-forme, θ agit sur un vecteur \vec{X} de $T_x M$ en donnant un réel, puisque θ et \vec{X} dépendent du point x avec régularité, ce nombre réel varie de façon régulière avec le point x , nous avons donc une fonction : $f : M \rightarrow \mathbb{R}$, telle que :

$$f : x \mapsto f(x) := \theta(x) \cdot \vec{X}(x).$$

Dans chaque espace cotangent nous pouvons définir une base "naturelle" donnée par les différentielles des fonctions définissant les cartes locales, si $x \in U_i \subset M$, alors une telle base est donnée par $d\varphi_1^{(i)}, \dots, d\varphi_m^{(i)}$. Car $\varphi(x)$ est une coordonnée sur \mathbb{R}^m , on utilise souvent la notation simplifiée : dx_1, \dots, dx_m "en oubliant" sur quelle carte on travaille. Dans cette base un élément $\theta(x) \in T_x^* M$ pourra s'écrire comme :

$$\theta(x) = \sum_{j=1}^m \theta_j(x) dx_j,$$

où la composante $\theta_j(x)$ est obtenue en appliquant la forme θ sur l'élément de base de $T_x M$, ∂_j .

⁴Une 1-forme est aussi appelée une *section* de $T^* M$, c'est-à-dire une application régulière $\theta : M \rightarrow T^* M$ telle que $\pi \circ \theta = id_M$, où $\pi : T^* M \rightarrow M$ est la projection naturelle définie par $\pi : (x, \theta) \in T_x^* M \mapsto \pi(x, \theta) = x$ et id_M est la fonction identité de M .

Comme nous l'avons vu pour les vecteurs, à chaque forme nous pouvons associer sa représentation locale, aussi dite *covecteur* pour la distinguer des vecteurs :

$$\theta(x) = \sum_{j=1}^m \theta_j(x) dx_j \Rightarrow \theta^{(j)}(x) = (\theta_1(x), \dots, \theta_m(x)),$$

où nous avons encore une fois abusé avec la notation en écrivant $\theta_j(x)$ au lieu de $\theta_j(\varphi_j(x))$. La fonction construite à partir d'une 1-forme et d'un vecteur aura une représentation locale :

$$f(x) = \theta(x) \cdot \vec{X}(x) = \sum_{j=1}^m \theta_j(x) \vec{X}_j(x).$$

Si $\theta^{(i)}$ et $\theta^{(j)}$ sont deux représentations locales de la même 1-forme, dans deux cartes différentes, alors ces objets sont reliés par la formule suivante :

$$(III.5.3) \quad \theta^{(j)}(\xi_j) = (A)^T \theta^{(i)}(\psi_{ij}(\xi_j)),$$

où encore un fois A est la matrice Jacobienne de ϕ_{ij} , fonction de transition entre les cartes (U_i, φ_i) et (U_j, φ_j) , et $\xi_j = \varphi_j(x)$ est l'image du point $x \in U_j \subset M$ dans la j -ème carte. Observez que les 1-formes ont une loi de transformation différente des vecteurs (cfr. (III.5.2)). Cette loi est compatible avec la loi de transformation de fonctions (III.5.1), car comme nous avons déjà dit avec une 1-forme, θ , et un vecteur \vec{X} on peut construire la fonction $f(x) = \theta(x)(\vec{X}(x))$; car les objets θ et \vec{X} sont globaux, nous voudrions la même propriété pour f , en effet si nous prenons deux représentatifs locaux pour la forme et le vecteur nous avons ⁵ :

$$f_i(\psi_{ij}(\xi_j)) = \theta^{(i)}(\psi_{ij}(\xi_j)) \cdot \vec{X}^{(i)}(\psi_{ij}(\xi_j)) = (A^{-1})^T \theta^{(j)}(\xi_j) \cdot A \vec{X}^{(j)}(\xi_j) = f_j(\xi_j),$$

Vice versa les relations (III.5.3) assurent qu'une famille de 1-forme admet une forme globale définie sur la variété.

L'exemple plus simple de 1-forme est la différentielle d'une fonction, $f : M \rightarrow \mathbb{R}$, la df est la 1-forme dont les composantes dans une carte locale donnée sont :

$$df(x) = \sum_{j=1}^m \frac{\partial f}{\partial x_j} dx_j.$$

5.4. 2-formes différentielles. Nous avons introduit les 1-formes comme applications linéaires définies sur un certain espace vectoriel, nous pouvons bien généraliser cette idée à des objets définis sur deux vecteurs à valeur toujours dans \mathbb{R} , c'est les *2-formes (externes)*. Une 2-forme est une application *bilinéaire, antisymétrique* qui à tout couple de éléments de $T_x M$ associe un réel : $\omega(x) : T_x M \times T_x M \rightarrow \mathbb{R}$. La propriété d'antisymétrie permet de dire que :

$$\omega(x)(\vec{X}, \vec{Y}) = -\omega(x)(\vec{Y}, \vec{X}) \quad \forall (\vec{X}, \vec{Y}) \in T_x M \times T_x M$$

En coordonnées locales, une 2-forme est représentée par une matrice, antisymétrique $(\omega_{ij})_{ij}$ (nous oublierons ici d'écrire les indices des cartes locales pour ne pas alourdir les notations) :

$$\omega(x)(\vec{X}, \vec{Y}) = \sum_{i,j=1}^m X_i(x) \omega_{ij}(x) Y_j(x);$$

⁵L'opération d'appliquer une 1-forme à un vecteur peut être vue comme un produit scalaire et donc nous avons la propriété suivante $B\theta \cdot \vec{X} = \theta \cdot (B)^T \vec{X}$ si nous avons des objets réels.

comme nous avons vu pour les 1-formes, les “composantes” de la 2-forme sont obtenues en l’appliquant aux vecteurs de la base de $T_x M$: $\omega_{ij}(x) = \omega(\partial_i, \partial_j)$.

Si $\omega^{(j)}$ et $\omega^{(i)}$ sont deux représentations locales de la même 2-forme, alors nous avons :

$$(III.5.4) \quad \omega^{(j)}(\xi_j) = (A)^T \omega^{(i)}(\psi_{ij}(\xi_j)) A,$$

avec les mêmes notations utilisées plus haut. Encore une fois ces relations permettent de définir un objet global en partant d’une famille d’objets locaux.

Une 2-forme peut être construite de deux façons en partant de 1-formes : *produit extérieur* et *différentiation extérieure*.

Soit données deux 1-formes, θ et θ' , alors leur produit extérieur est défini comme :

$$(III.5.5) \quad (\theta \wedge \theta')(\vec{X}, \vec{Y}) := (\theta \cdot \vec{X})(\theta' \cdot \vec{Y}) - (\theta \cdot \vec{Y})(\theta' \cdot \vec{X}),$$

on peut vérifier que l’objet ainsi défini est bien une 2-forme.

Dans chaque carte locale nous avons la base des 1-formes, dx_1, \dots, dx_m , avec le produit extérieur nous pouvons définir une base pour les 2-formes :

$$dx_i \wedge dx_j \quad 1 \leq i, j \leq m,$$

on peut aisément vérifier que seulement $m(m-1)/2$ d’entre eux sont indépendants (utiliser l’antisymétrie pour mettre à zéro ceux avec les mêmes indices et pour relier ceux avec les indices échangés). Dans cette base la représentation locale d’une 2-forme est :

$$\omega(x) = \sum_{1 \leq i, j \leq m} \omega_{ij}(x) dx_i \wedge dx_j.$$

Une deuxième façon de “construire” une 2-forme en partant d’une 1-forme, est en procédant par différentiation extérieure ; si dans une carte locale, la 1-forme admet l’écriture $\theta(x) = \sum_{j=1}^m \theta_j(x) dx_j$, alors par définition :

$$(III.5.6) \quad d\theta := \sum_{i=1}^m d\theta_i \wedge dx_i = \sum_{i,j=1, i < j}^m \left(\frac{\partial \theta_i}{\partial x_j} - \frac{\partial \theta_j}{\partial x_i} \right) dx_i \wedge dx_j.$$

On peut vérifier que cela définit une 2-forme locale et de plus les relations (III.5.4) s’appliquent de façon à définir un objet global.

5.5. Quelques définitions. Une 1-forme est dite *fermée* si sa dérivée extérieure vaut zéro : $d\theta = 0$. Une 1-forme est dite *exacte* s’il existe une fonction f , telle que $\theta = df$. Toute forme exacte est aussi fermée : $d(df) = 0$, mais le contraire n’est pas vrai, il s’avère que la fonction qui donne la forme exacte n’est définie que localement, notons aussi que pour les ouverts simplement connexes de \mathbb{R}^m les deux notions coïncident (Lemme de Poincaré).

Une 2-forme est dite *non-dégénérée* si :

$$\omega(\vec{X}, \vec{Y}) = 0 \quad \forall \vec{Y} \Rightarrow \vec{X} = 0,$$

dans une représentation locale cela se réécrit comme $\det \omega_{ij} \neq 0$. Toute 2-forme non-dégénérée permet d’établir une correspondance biunivoque entre 1-formes et vecteurs : pour tout champ de vecteurs \vec{Y} est bien définie la 1-forme $\theta = \omega(\cdot, \vec{Y})$, c’est-à-dire la 1-forme telle que sur un vecteur \vec{X} elle donne : $\theta \cdot \vec{X} = \omega(\vec{X}, \vec{Y})$. Vice

versa si nous avons une 1-forme θ , alors grâce à la condition de non-dégénération il existe un seul vecteur ⁶ \vec{Y} , tel que $\theta \cdot \vec{X} = \omega(\vec{X}, \vec{Y})$ pour tout \vec{X} .

Si une 2-forme est antisymétrique et non-dégénérée alors la dimension de la variété est forcément paire, en effet :

$$\det \omega = \det(\omega)^T = \det(-\omega) = (-1)^m \det \omega,$$

et puisque $\det \omega \neq 0$ il suit nécessairement que $m = 2n$.

6. Condition de Lie

Le but de cette section est de caractériser les transformations canoniques à l'aide des formes différentielles ; cette nouvelle caractérisation sera utilisée dans la section suivante pour “construire” des transformations canoniques.

Dans cette section nous utilisons le formalisme des formes différentielles, vu dans le cours de géométrie et présenté brièvement dans la section précédente, initialement nous travaillerons dans \mathbb{R}^m , dans une autre section nous aborderons brièvement le problème sur les variétés. Nous commençons par donner une définition.

DÉFINITION 6.1. *Une forme différentielle θ dans \mathbb{R}^{2m+1} :*

$$(III.6.1) \quad \theta = \sum_{i=1}^{2m+1} \theta_i(\vec{x}) dx_i,$$

est dite non singulière si la matrice $A(\vec{x})$, antisymétrique $(2m+1) \times (2m+1)$, définie par ⁷ :

$$(III.6.2) \quad A_{ij} = \frac{\partial \theta_i}{\partial x_j} - \frac{\partial \theta_j}{\partial x_i},$$

est de rang maximal (c'est-à-dire $2m$). Le noyau de $A(\vec{x})$, $\ker(A) = \{\vec{v} \in \mathbb{R}^{2m+1} : A(\vec{x})\vec{v} = 0\}$ détermine, quand \vec{x} varie, un champ de directions appelées directions de rotationnel. Les courbes intégrales du champ de directions de rotationnel sont appelées courbes de rotationnel.

EXEMPLE 6.2. *La forme $\theta = x_2 dx_1 + x_3 dx_2 + x_1 dx_3$ dans \mathbb{R}^3 est non singulière. En effet la matrice A associée est :*

$$A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix},$$

qui est de rang 2. Le noyau de A est formé des vecteurs $\vec{v} \in \mathbb{R}^3$ tels que :

$$A\vec{v} = 0 \Rightarrow v_2 = v_3, v_1 = v_3 \quad \text{et} \quad v_1 = v_2,$$

c'est-à-dire $\vec{v} = a(1, 1, 1)^T$, pour $a \in \mathbb{R}$.

EXEMPLE 6.3. *La forme $\theta = x_1 dx_2 + \frac{1}{2}(x_1^2 + x_2^2) dx_3$ dans \mathbb{R}^3 est non singulière. En effet la matrice A associée est :*

$$A = \begin{pmatrix} 0 & -1 & -x_1 \\ 1 & 0 & -x_2 \\ x_1 & x_2 & 0 \end{pmatrix},$$

⁶S'il existait $\vec{Y}_1 \neq \vec{Y}_2$ tels que $\theta \cdot \vec{X} = \omega(\vec{X}, \vec{Y}_1) = \omega(\vec{X}, \vec{Y}_2)$ pour tout \vec{X} alors $0 = \omega(\vec{X}, \vec{Y}_1) - \omega(\vec{X}, \vec{Y}_2) = \omega(\vec{X}, \vec{Y}_1 - \vec{Y}_2)$ pour tout \vec{X} , mais alors la condition de non-dégénération nous donnerait $\vec{Y}_1 - \vec{Y}_2 = 0$ contre l'hypothèse.

⁷Observez que (A_{ij}) est la matrice représentative de la forme $d\theta$.

qui est de rang 2. Le noyau de A est formé des vecteurs $\vec{v} \in \mathbb{R}^3$ tels que :

$$A\vec{v} = 0 \Rightarrow v_2 = -x_1v_3, v_1 = x_2v_3 \quad \text{et} \quad x_1v_1 = -x_2v_2,$$

c'est-à-dire $\vec{v} = (x_2, -x_1, 1)$.

Considérons maintenant une courbe fermée γ , alors les lignes de rotationnels passant par la courbe γ définissent une surface dans \mathbb{R}^{2m+1} et le *Lemme de Stokes* permet d'identifier les intégrales de la forme non-singulière sur les deux courbes (voir Fig. 5).

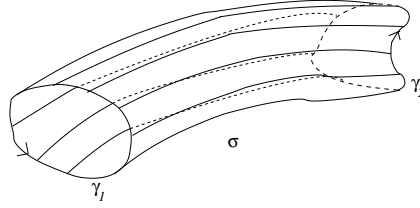


FIG. 5. Géométrie pour le Lemme de Stokes.

LEMME 6.4 (Lemme de Stokes). Soit θ une 1-forme non-singulière dans \mathbb{R}^{2m+1} et soit γ_1 et γ_2 deux courbes fermées appartenant à la même surface du rotationnel de θ alors :

$$\int_{\gamma_1} \theta = \int_{\gamma_2} \theta.$$

DÉMONSTRATION. La démonstration est basée sur le Théorème de Stokes⁸ et sur le fait que $d\theta = 0$ sur la surface du rotationnel grâce à la condition de non-singularité, par définition (III.5.6) :

$$d\theta = \sum_{i,j=1}^{2m+1} \left(\frac{\partial \theta_i}{\partial x_j} - \frac{\partial \theta_j}{\partial x_i} \right) dx_i \wedge dx_j,$$

et donc pour tout $\vec{v} \in \ker(A)$ nous avons $d\theta \cdot \vec{v} = 0$.

Finalement en prenant comme variété σ la surface du rotationnel pour le Théorème de Stokes nous avons :

$$\int_{\sigma} d\theta = \int_{\partial\sigma} \theta,$$

mais $\partial\sigma = \gamma_1 \cup \gamma_2^{-1}$ (c'est-à-dire γ_2 parcourue en direction inverse), donc :

$$0 = \int_{\sigma} d\theta = \int_{\partial\sigma} \theta = \int_{\gamma_1} \theta - \int_{\gamma_2} \theta.$$

□

⁸*Théorème de Stokes* Soit M une variété 2-dimensionnelle avec bord et orientée, alors $\int_M d\theta = \int_{\partial M} \theta$, où θ est une 1-forme sur M .

Le Théorème est vrai pour une variété de dimension quelconque, l , et une $(l-1)$ -forme. La formule d'intégration vue dans le cours de première année est donc un cas particulier avec $M = [a, b] \subset \mathbb{R}$, θ une 0-forme, c'est-à-dire une fonction, et donc : $\int_a^b f' dx = \int_{[a,b]} d\theta = \int_{\partial[a,b]} \theta = f \Big|_a^b = f(b) - f(a)$.

Nous pouvons maintenant introduire la *forme de Poincaré–Cartan* qui va nous permettre de caractériser les transformations canoniques.

THÉOREME 6.5. *La 1-forme différentielle :*

$$(III.6.3) \quad \theta = \sum_{j=1}^m p_j dq_j - \mathcal{H}(\vec{p}, \vec{q}, t) dt,$$

en \mathbb{R}^{2m+1} est non-singulière et elle est dite *forme de Poincaré–Cartan*. Ses lignes de rotationnel sont les courbes intégrales des équations d'Hamilton associées à l'hamiltonien \mathcal{H} .

DÉMONSTRATION. La matrice associée à la forme θ est donnée par :

$$A(\vec{p}, \vec{q}, t) = \begin{pmatrix} \emptyset & -\mathbb{I} & \nabla_{\vec{p}} \mathcal{H} \\ \mathbb{I} & \emptyset & \nabla_{\vec{q}} \mathcal{H} \\ -(\nabla_{\vec{p}} \mathcal{H})^T & -(\nabla_{\vec{q}} \mathcal{H})^T & 0 \end{pmatrix}.$$

Évidemment le rang de A est $2m$ pour tout (\vec{p}, \vec{q}, t) car elle contient comme mineur la matrice \mathcal{J} . En outre le vecteur $\vec{v}(\vec{p}, \vec{q}, t) = (-\nabla_{\vec{q}} \mathcal{H}, \nabla_{\vec{p}} \mathcal{H}, 1)^T$ appartient au noyau de A :

$$A\vec{v} = \begin{pmatrix} \emptyset & -\mathbb{I} & \nabla_{\vec{p}} \mathcal{H} \\ \mathbb{I} & \emptyset & \nabla_{\vec{q}} \mathcal{H} \\ -(\nabla_{\vec{p}} \mathcal{H})^T & -(\nabla_{\vec{q}} \mathcal{H})^T & 0 \end{pmatrix} \begin{pmatrix} -\nabla_{\vec{q}} \mathcal{H} \\ \nabla_{\vec{p}} \mathcal{H} \\ 1 \end{pmatrix} = 0,$$

et donc les lignes de rotationnel de θ coïncident avec les solutions des équations d'Hamilton dans l'espace des phases *étendu*, c'est-à-dire en incluant aussi le temps. \square

Le Lemme de Stokes pour la forme de Poincaré–Cartan (connu en littérature comme Théorème de l'*invariant intégrale de Poincaré–Cartan*) permet de caractériser les transformations canoniques en termes de géométrie de l'espace des phases. En effet les transformations canoniques envoient un système d'équations d'Hamilton pour un hamiltonien \mathcal{H} dans un nouveau système d'équations d'Hamilton pour un nouveau hamiltonien \mathcal{K} , et donc elles envoient les surfaces de rotationnel d'un système sur l'autre.

PROPOSITION 6.6. *Une transformation canonique envoie les surfaces de rotationnel de la forme de Poincaré–Cartan (III.6.3) dans les surfaces de rotationnel de la forme :*

$$(III.6.4) \quad \Theta = \sum_{j=1}^m P_j dQ_j - \mathcal{K}(\vec{P}, \vec{Q}, t) dt.$$

EXEMPLE 6.7. *Considérons la transformation :*

$$p = \alpha\sqrt{P}\cos(\gamma Q), \quad q = \beta\sqrt{P}\sin(\gamma Q) \quad \alpha\beta\gamma \neq 0.$$

Déterminer pour quelles valeurs des paramètres elle est canonique et ensuite calculer la forme de Poincaré–Cartan dans les deux systèmes des coordonnées.

Nous devons déterminer les paramètres tels que la matrice Jacobienne de la transformation soit symplectique pour tout (p, q, t) :

$$A = \begin{pmatrix} \frac{\alpha}{2\sqrt{P}} \cos(\gamma Q) & -\gamma\alpha\sqrt{P} \sin(\gamma Q) \\ \frac{\beta}{2\sqrt{P}} \sin(\gamma Q) & \gamma\beta\sqrt{P} \cos(\gamma Q) \end{pmatrix},$$

or nous avons vu que $Sp(1, \mathbb{R}) \sim SL(2, \mathbb{R})$, donc il est suffisant de calculer le déterminant de A et de l'imposer égal à 1 :

$$1 = \det A = \frac{\alpha\beta\gamma}{2}.$$

Soit donc $\alpha\beta\gamma = 2$, et calculons la différence des formes de Poincaré–Cartan dans les deux systèmes de coordonnées exprimée dans les nouvelles coordonnées :

$$\begin{aligned} \theta - \Theta &= pdq - \mathcal{H}dt - (PdQ - \mathcal{K}dt) \\ &= \alpha\sqrt{P}\cos(\gamma Q) \left(\frac{\beta}{2\sqrt{P}}\sin(\gamma Q)dP + \beta\gamma\sqrt{P}\cos(\gamma Q)dQ \right) - \mathcal{H}dt - (PdQ - \mathcal{K}dt), \end{aligned}$$

la transformation ne dépend pas du temps de façon explicite et donc les deux hamiltoniens sont liés par la relation $\mathcal{H} = \mathcal{K}$, donc :

$$\begin{aligned} \theta - \Theta &= \frac{\alpha\beta}{2}\cos(\gamma Q)\sin(\gamma Q)dP + \alpha\beta\gamma P\cos^2(\gamma Q)dQ - PdQ \\ &= \frac{1}{2\gamma}\sin(2\gamma Q)dP - 2P\left(\frac{1}{2} - \cos^2(\gamma Q)\right)dQ \\ &= \frac{1}{2\gamma}\sin(2\gamma Q)dP + P\cos(2\gamma Q)dQ. \end{aligned}$$

Observons que le dernier terme est la différentielle d'une fonction (la dernière écriture de la 1-forme nous donne une forme exacte) :

$$d\left(\frac{P}{2\gamma}\sin(2\gamma Q)\right) = \frac{1}{2\gamma}\sin(2\gamma Q)dP + P\cos(2\gamma Q)dQ,$$

donc nous pouvons conclure que :

$$\theta - \Theta = d\left(\frac{P}{2\gamma}\sin(2\gamma Q)\right),$$

or deux formes qui diffèrent pour une forme exacte ont la même surface de rotationnel car $d\left[d\left(\frac{P}{2\gamma}\sin(2\gamma Q)\right)\right] = 0$ et donc $d\theta = d\Theta$.

Nous allons maintenant montrer que non seulement le fait d'envoyer les surfaces de rotationnel d'un système sur celles d'un deuxième est une condition nécessaire pour avoir une transformation canonique, mais cette condition s'avère aussi suffisante.

THÉORÈME 6.8 (Condition de Lie). Une transformation de coordonnées $p_i = p_i(\vec{P}, \vec{Q}, t)$, $q_i = q_i(\vec{P}, \vec{Q}, t)$, pour $i = 1, \dots, m$ est canonique si et seulement si la différence entre les formes de Poincaré–Cartan écrite dans les deux systèmes de coordonnées est une forme exacte, c'est-à-dire il existe une fonction régulière f telle que :

$$(III.6.5) \quad \tilde{\theta} - \tilde{\Theta} = \sum_{j=1}^m \left(p_j \tilde{d}q_j - P_j \tilde{d}Q_j \right) = \tilde{d}f,$$

où nous avons introduit la différentiation extérieure à temps fixé, c'est-à-dire pour toute fonction $g(\vec{P}, \vec{Q}, t)$ régulière, alors $\tilde{d}g := dg - \frac{\partial g}{\partial t}dt = \sum_{j=1}^m \left(\frac{\partial g}{\partial P_j}dP_j + \frac{\partial g}{\partial Q_j}dQ_j \right)$.

REMARQUE 6.9. Si la transformation est complètement canonique, c'est-à-dire elle ne dépend pas de façon explicite du temps, alors $\tilde{d} = d$ et la fonction f du Théorème peut être choisie elle aussi indépendante du temps.

EXEMPLE 6.10. Dans l'exemple 4.5 nous avons montré que la transformation de variables induite d'une transformation sur les seules variables \vec{q} est canonique. Avec le Théorème précédent la démonstration devient triviale.

Soit donc $\vec{Q} = \vec{Q}(\vec{t})$ et $\vec{P} = ((A)^T)^{-1}\vec{p}$, où A est la matrice Jacobienne du changement de variables, $A = \frac{\partial Q_i}{\partial q_j}$; alors la différence entre les formes de Poincaré–Cartan s'écrit (notation indices répétés égal indices additionnés) :

$$p_j dq_j - P_j dQ_j = p_j dq_j - p_l (A^{-1})_{lj} A_{jk} dq_k = p_j dq_j - p_l \delta_{lk} dq_k = p_j dq_j - p_k dq_k \equiv 0.$$

Nous concluons cette section avec le résultat suivant qui caractérise les transformations canoniques en termes de forme de Poincaré–Cartan :

THÉORÈME 6.11. Si la transformation de coordonnées $p_i = p_i(\vec{P}, \vec{Q}, t)$, $q_i = q_i(\vec{P}, \vec{Q}, t)$, pour $i = 1, \dots, m$ est canonique, et si nous dénotons la nouvelle forme de Poincaré–Cartan par :

$$(III.6.6) \quad \Theta = \sum_{j=1}^m P_j dQ_j - \mathcal{K}(\vec{P}, \vec{Q}, t) dt,$$

alors il existe une fonction régulière $f(\vec{P}, \vec{Q}, t)$ telle que :

$$(III.6.7) \quad \sum_{j=1}^m (p_j dq_j - P_j dQ_j) + (\mathcal{K} - \mathcal{H}) dt = df,$$

c'est-à-dire la différence entre les deux formes de Poincaré–Cartan est une forme exacte.

Vice versa si la transformation de coordonnées $p_i = p_i(\vec{P}, \vec{Q}, t)$, $q_i = q_i(\vec{P}, \vec{Q}, t)$, pour $i = 1, \dots, m$ est telle qu'il existe deux fonctions $\mathcal{K}(\vec{P}, \vec{Q}, t)$ et $f(\vec{P}, \vec{Q}, t)$, telles que Θ définie comme dans (III.6.6) vérifie la relation (III.6.7), alors la transformation est canonique et le nouvel hamiltonien est \mathcal{K} .

7. Fonctions génératrices (de transformations canoniques)

Dans la section précédente nous avons caractérisé les transformations canoniques, nous allons maintenant nous intéresser à une procédure explicite qui nous permettra de construire ces transformations canoniques.

Nous avons vu que la *condition de Lie* (cfr. (III.6.5) ou (III.6.7)) est une condition nécessaire et suffisante pour avoir une transformation canonique.

Supposons qu'on ait deux groupes de coordonnées, les vieilles (\vec{p}, \vec{q}) et les nouvelles (\vec{P}, \vec{Q}) , et supposons que l'on connaisse la relation qui les relie :

$$(III.7.1) \quad \vec{P} = \vec{P}(\vec{q}, \vec{Q}, t) \quad \text{et} \quad \vec{p} = \vec{p}(\vec{q}, \vec{Q}, t),$$

observons que nous n'avons pas un changement de coordonnées car nous savons exprimer le groupes de vieilles variables, \vec{p} , et un groupe des nouvelles, \vec{P} , en fonction des restantes (\vec{q}, \vec{Q}) . Nous nous demandons sous quelle condition nous pouvons obtenir une transformation canonique (c'est-à-dire exprimer les nouvelles en termes des vieilles et préserver la structure des équations d'Hamilton).

Supposons qu'on ait résolu pour le moment le problème de savoir exprimer les nouvelles variables en fonction des vieilles, alors la condition de Lie (III.6.7) nous

assure l'existence d'une transformation canonique si nous trouvons les fonctions \mathcal{K} et F ; la condition de Lie se réécrit dans le cas présent comme :

(III.7.2)

$$\sum_{j=1}^m p_j(\vec{q}, \vec{Q}, t) dq_j - \mathcal{H}(\vec{p}(\vec{q}, \vec{Q}, t), \vec{q}, t) dt - \left(\sum_{j=1}^m P_j(\vec{q}, \vec{Q}, t) dQ_j - \mathcal{K}(\vec{P}(\vec{q}, \vec{Q}, t), \vec{Q}, t) dt \right) = dF(\vec{q}, \vec{Q}, t).$$

Mais la différentielle de la fonction F est donnée par (rappelons que les variables \vec{q} et \vec{Q} sont supposées être indépendantes) :

$$(III.7.3) \quad dF(\vec{q}, \vec{Q}, t) = \sum_{j=1}^m \left(\frac{\partial F}{\partial q_j} dq_j + \frac{\partial F}{\partial Q_j} dQ_j \right) + \frac{\partial F}{\partial t} dt,$$

et pour satisfaire les conditions (III.7.2) et (III.7.3) il faut que les relations suivantes soient vérifiées :

$$(III.7.4) \quad p_j = \frac{\partial F}{\partial q_j}, \quad P_j = -\frac{\partial F}{\partial Q_j} \quad \text{et} \quad \mathcal{K} = \mathcal{H} + \frac{\partial F}{\partial t} \quad \forall j = 1, \dots, m.$$

Il nous reste le problème d'exprimer les nouvelles coordonnées en fonction des vieilles, mais pour cela (Théorème de la fonction implicite) il est suffisant de demander :

$$(III.7.5) \quad \det \frac{\partial p_i}{\partial Q_j} \neq 0,$$

car de cette façon nous pouvons inverser la seconde relation des (III.7.1) pour exprimer :

$$\vec{Q} = \vec{Q}(\vec{q}, \vec{p}, t),$$

et substituer dans la première relation de (III.7.1). Les variables \vec{P} sont ainsi obtenues :

$$\vec{P} = \hat{\vec{P}}(\vec{q}, \vec{Q}(\vec{q}, \vec{p}, t), t) = \vec{P}(\vec{p}, \vec{q}, t);$$

la relation (III.7.5) est équivalente grâce à (III.7.4) à la relation :

$$(III.7.6) \quad \det \frac{\partial^2 F}{\partial q_i \partial Q_j} \neq 0.$$

Nous pouvons résumer ce que nous avons fait dans la définition suivante.

DÉFINITION 7.1 (Fonction génératrice (du premier type)). *Une fonction $F(\vec{q}, \vec{Q}, t)$ qui vérifie la relation (III.7.6) induit une transformation canonique sous les conditions (III.7.4). Cette fonction est dite fonction génératrice (du premier type).*

EXEMPLE 7.2. *La fonction $F(q, Q) = \frac{m\omega}{2} q^2 \operatorname{ctg} Q$ détermine une transformation canonique, de plus elle permet de transformer l'hamiltonien $\mathcal{H}(p, q) = \frac{p^2}{2m} + \frac{m\omega^2 q^2}{2}$ dans le nouveau hamiltonien $\mathcal{K}(P, Q) = \omega P$.*

La condition (III.7.6) est vérifiée car :

$$\frac{\partial^2 F}{\partial q \partial Q} = -m\omega q(1 + \operatorname{ctg}^2 Q) \neq 0 \quad \text{si } q \neq 0 \quad (\text{il faut aussi exclure } Q = 0 \text{ pour avoir la régularité});$$

la transformation induite est (cfr. (III.7.4)) :

$$p = m\omega q \operatorname{ctg} Q \quad \text{et} \quad P = \frac{m\omega}{2} q^2 (1 + \operatorname{ctg}^2 Q),$$

il nous reste maintenant à exprimer un groupe de variables en fonction des autres, par exemple (p, q) en fonction des nouvelles :

$$q = \sqrt{\frac{2P}{m\omega(1 + \cot^2 Q)}} = \sqrt{\frac{2P}{m\omega}} \sin Q,$$

et donc :

$$p = m\omega \sqrt{\frac{2P}{m\omega}} \cos Q = \sqrt{2Pm\omega} \cos Q.$$

Le nouveau hamiltonien est le transformé du vieux, car la transformation ne dépend pas du temps de façon explicite :

$$\begin{aligned} \mathcal{K}(P, Q) &= \mathcal{H}(p(P, Q), q(P, Q)) = \frac{1}{2m} \left(\sqrt{2Pm\omega} \cos Q \right)^2 + \frac{m\omega^2}{2} \left(\sqrt{\frac{2P}{m\omega}} \sin Q \right)^2 \\ &= P\omega. \end{aligned}$$

Dans les nouvelles variables la dynamique est triviale :

$$\begin{cases} \dot{P} &= -\frac{\partial \mathcal{K}}{\partial Q} = 0 \\ \dot{Q} &= \frac{\partial \mathcal{K}}{\partial P} = \omega \end{cases}$$

donc

$$P(t) = P(0) \quad \text{et} \quad Q(t) = Q(0) + \omega t,$$

et finalement dans les variables de départ nous avons :

$$p(t) = \sqrt{2P(0)m\omega} \cos(\omega t + Q(0)) \quad \text{et} \quad q(t) = \sqrt{\frac{2P(0)}{m\omega}} \sin(\omega t + Q(0)).$$

La fonction génératrice que nous avons introduite dépend des variables (q, Q) , mais nous aurions pu prendre n'importe quel autre ensemble de vieilles et nouvelles variables (sans répétition) ; nous avons aussi vu que la transformation qui échange les variables q avec les variables p (à un signe près) est canonique et donc quitte à renommer les variables nous pouvons introduire quatre classes de fonctions génératrices :

(III.7.7)

$$\begin{cases} \text{(premier type)} & F_1(\vec{q}, \vec{Q}, t) & \text{avec } \det \left(\frac{\partial^2 F_1}{\partial q_i \partial Q_j} \right) \neq 0 & \text{et } p_i = \partial_{q_i} F_1 \quad P_i = -\partial_{Q_i} F_1 \\ \text{(deuxième type)} & F_2(\vec{q}, \vec{P}, t) & \text{avec } \det \left(\frac{\partial^2 F_2}{\partial q_i \partial P_j} \right) \neq 0 & \text{et } p_i = \partial_{q_i} F_2 \quad Q_i = \partial_{P_i} F_2 \\ \text{(troisième type)} & F_3(\vec{p}, \vec{Q}, t) & \text{avec } \det \left(\frac{\partial^2 F_3}{\partial Q_i \partial p_j} \right) \neq 0 & \text{et } q_i = -\partial_{Q_i} F_3 \quad P_i = -\partial_{Q_i} F_3 \\ \text{(quatrième type)} & F_4(\vec{p}, \vec{P}, t) & \text{avec } \det \left(\frac{\partial^2 F_4}{\partial P_i \partial p_j} \right) \neq 0 & \text{et } q_i = -\partial_{Q_i} F_4 \quad Q_i = \partial_{P_i} F_4. \end{cases}$$

REMARQUE 7.3 (Comment passer d'une version aux autres). Si une transformation admet plus d'une fonction génératrice ces fonctions sont alors liées par des relations.

Par exemple pour lier F_2 avec F_1 , nous écrivons la condition de Lie pour F_1 (notation indices répétés indique sommation) :

$$(III.7.8) \quad p_i dq_i - \mathcal{H} dt - (P_i dQ_i - \mathcal{K} dt) = dF_1,$$

mais F_2 est une fonction de \vec{q} et \vec{P} , donc :

$$dF_2 = \frac{\partial F_2}{\partial q_i} dq_i + \frac{\partial F_2}{\partial P_i} dP_i + \frac{\partial F_2}{\partial t} dt.$$

Il faut modifier (III.7.8) pour faire apparaître la différentielle d'une fonction de \vec{q} et \vec{P} :

$$dF_1 = p_i dq_i - \mathcal{H} dt - (P_i dQ_i - \mathcal{K} dt) - Q_i dP_i + Q_i dP_i = p_i dq_i - \mathcal{H} dt + \mathcal{K} dt - d(P_i Q_i) + Q_i dP_i,$$

donc nous obtenons

$$F_2 = F_1 + P_i Q_i,$$

car le terme introduit $d(P_i Q_i)$ vérifie $d^2(P_i Q_i) = 0$. Par identification des différentielles, nous pouvons conclure :

$$p_i = \frac{\partial F_2}{\partial q_i}, Q_i = \frac{\partial F_2}{\partial Q_i} \text{ et } \mathcal{K} - \mathcal{H} = \frac{\partial F_2}{\partial t}.$$

De façon similaire pour les autres :

$$F_3 = F_1 - p_i q_i \text{ et } F_4 = F_3 + P_i Q_i.$$

REMARQUE 7.4. Les transformations définies initialement sur les variables \vec{q} , $\vec{Q} = \vec{Q}(\vec{q}, t)$, appartiennent au deuxième type :

$$F_2(\vec{q}, \vec{P}, t) = P_i Q_i(\vec{q}, t),$$

d'où :

$$p_i = P_j \frac{\partial Q_j}{\partial q_i} = (A_{ij})^T P_j,$$

avec A matrice Jacobienne.

La transformation identique rentre aussi dans cette classe :

$$F_2(\vec{q}, \vec{P}, t) = P_i q_i,$$

et donc

$$p_i = \frac{\partial F_2}{\partial q_i} = P_i \text{ et } Q_i = \frac{\partial F_2}{\partial P_i} = q_i.$$

REMARQUE 7.5. La transformation qui échange (à un signe près) les variables est dans la quatrième classe :

$$F_4(\vec{p}, \vec{P}, t) = P_i p_i,$$

et

$$q_i = -\frac{\partial F_4}{\partial p_i} = -P_i \text{ et } Q_i = \frac{\partial F_4}{\partial P_i} = p_i.$$

8. Parenthèses de Poisson

Nous avons déjà remarqué l'analogie entre les matrices symplectiques et les matrices orthogonales (il suffit de remplacer dans leur définition la matrice \mathcal{J} avec la matrice identité). Cela permet d'introduire un *produit symplectique*, au lieu d'un produit scalaire :

$$(III.8.1) \quad (\vec{x})^T * \vec{y} := (\vec{x})^T \mathcal{J} \vec{y} = x_i \mathcal{J}_{ij} y_j \quad \forall \vec{x}, \vec{y} \in \mathbb{R}^{2m}.$$

Avec cette notion nous pouvons définir la *parenthèse de Poisson* entre deux fonctions régulières $f, g : \mathbb{R}^{2m+1} \rightarrow \mathbb{R}$, comme le produit symplectique de leur gradient :

$$(III.8.2) \quad \{f, g\} := (\nabla f)^T * \nabla g = (\nabla f)^T \mathcal{J} \nabla g;$$

Si $\vec{x} = (\vec{p}, \vec{q})$ alors la parenthèse de Poisson s'écrit comme

$$(III.8.3) \quad \{f, g\} = \sum_{i=1}^m \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i} \frac{\partial f}{\partial p_i} \right).$$

La parenthèse de Poisson est un instrument fondamental dans la mécanique hamiltonienne, nous montrerons son usage pour la recherche des intégrales premières et pour la caractérisation des transformations canoniques. Avant de faire cela nous proposons deux remarques.

REMARQUE 8.1. *Si nous calculons les parenthèses de Poisson avec les fonctions p_j et \mathcal{H} ou bien avec q_j et \mathcal{H} alors nous trouvons :*

$$\{p_j, \mathcal{H}\} = \sum_{i=1}^m \left(\frac{\partial p_j}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \mathcal{H}}{\partial q_i} \frac{\partial p_j}{\partial p_i} \right) = \sum_{i=1}^m \left(0 \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \mathcal{H}}{\partial q_i} \delta_{ij} \right) = - \frac{\partial \mathcal{H}}{\partial q_j} = \dot{p}_j,$$

et

$$\{q_j, \mathcal{H}\} = \dot{q}_j.$$

Donc les équations d'Hamilton peuvent être réécrites en termes de parenthèses de Poisson.

REMARQUE 8.2 (Parenthèses de Poisson fondamentales). *On appelle parenthèses de Poisson fondamentales les parenthèses calculées avec les variables p_i et q_i :*

$$\{p_i, p_j\} = \{q_i, q_j\} = 0 \quad \text{et} \quad \{p_i, q_j\} = -\delta_{ij}.$$

La proposition suivante réunit quelques propriétés des parenthèses de Poisson.

PROPOSITION 8.3. *Les parenthèses de Poisson vérifient :*

- (1) *Elles définissent une forme bilinéaire antisymétrique sur les fonctions, c'est-à-dire, l'application : $(f, g) \mapsto \{f, g\}$ est linéaire dans les deux "variables" f et g et $\{f, g\} = -\{g, f\}$;*

- (2) *Règle de Leibniz :*

$$\{f_1 f_2, g\} = f_1 \{f_2, g\} + f_2 \{f_1, g\};$$

- (3) *Identité de Jacobi :*

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0;$$

- (4) *Si on introduit l'opérateur $D_f = \{f, \cdot\}$, alors*

$$[D_f, D_g] = D_{\{f, g\}};$$

- (5) *Non-dégénération, si $\vec{x}_0 \in \mathbb{R}^{2m}$ n'est pas un point critique de f , alors il existe une fonction g telle que $\{f, g\}(\vec{x}_0) \neq 0$.*

8.1. Parenthèses de Poisson et intégrales premières. Soit $\mathcal{H} : \mathbb{R}^{2m+1} \rightarrow \mathbb{R}$ un hamiltonien et soit

$$(III.8.4) \quad \dot{\vec{x}} = \mathcal{J} \nabla_{\vec{x}} \mathcal{H},$$

les équations d'Hamilton associées. Supposons que pour toute donnée initiale $\vec{x}(0)$, le système (III.8.4) ait une solution qui existe pour tout t , dans ce cas le flot hamiltonien, Φ^t , définit un opérateur U^t qui, à toute fonction définie sur l'espace des phases, associe la valeur de cette fonction sur la solution au temps t des équations d'Hamilton :

$$(U^t f)(\vec{x}(0), 0) = f(\Phi^t(\vec{x}(0)), t) \equiv f(\vec{x}(t), t).$$

Une fonction $f : \mathbb{R}^{2m+1} \rightarrow \mathbb{R}$ est appelée *intégrale première* (du flot Hamiltonien) si et seulement si, quel que soit $\vec{x}_0 \in \mathbb{R}^{2m}$ et $t \in \mathbb{R}$ on a :

$$(III.8.5) \quad f(\Phi^t(\vec{x}_0), t) = f(\vec{x}_0, t).$$

C'est-à-dire la fonction f est constante sur les orbites du flot hamiltonien.

Puisque la dérivée de f calculée le long du flot est donnée par :

$$\left. \frac{df}{dt} \right|_{\text{flot}} = \frac{\partial f}{\partial t} + (\nabla_{\vec{x}} f)^T \dot{\vec{x}} = \frac{\partial f}{\partial t} + (\nabla_{\vec{x}} f)^T \mathcal{J} \nabla_{\vec{x}} \mathcal{H},$$

alors avec la définition de parenthèse de Poisson nous avons :

$$\left. \frac{df}{dt} \right|_{\text{flot}} = \frac{\partial f}{\partial t} + \{f, \mathcal{H}\}.$$

Nous pouvons finalement énoncer le résultat suivant

THÉORÈME 8.4. *Une fonction $f : \mathbb{R}^{2m} \rightarrow \mathbb{R}$, indépendante du temps, est une intégrale première du flot hamiltonien si et seulement si*

$$\{f, \mathcal{H}\} = 0.$$

8.2. Parenthèses de Poisson et transformations canoniques. La définition que nous avons donnée de parenthèse de Poisson est liée au système de coordonnées choisies, par contre le fait d'être une intégrale première est indépendant du choix des coordonnées et en plus est invariant par transformation canonique. Nous devons donc étudier comment les parenthèses de Poisson se transforment par un changement de coordonnées.

THÉORÈME 8.5. *Les faits suivants sont équivalents :*

- (1) *la transformation $\vec{x} = \vec{x}(\vec{X}, t)$ est canonique ;*
- (2) *pour tout couple de fonctions régulières $f(\vec{x}, t)$ et $g(\vec{x}, t)$, si nous appelons $F(\vec{X}, t) = f(\vec{x}(\vec{X}, t), t)$ et $G(\vec{X}, t) = g(\vec{x}(\vec{X}, t), t)$ leurs transformées dans les nouvelles coordonnées, alors :*

$$(III.8.6) \quad \{f, g\}_{\vec{x}} = \{F, G\}_{\vec{X}},$$

pour tout instant t , où nous avons dénoté par $\{f, g\}_{\vec{x}}$ les parenthèses de Poisson dans les coordonnées $\vec{x} = (\vec{p}, \vec{q})$ et $\{F, G\}_{\vec{X}}$ dans les nouvelles $\vec{X} = (\vec{P}, \vec{Q})$;

- (3) *Pour tout $i, j \in \{1, \dots, m\}$ et pour tout instant t nous avons :*

$$(III.8.7) \quad \{P_i, P_j\}_{\vec{x}} = \{Q_i, Q_j\}_{\vec{x}} = 0 \quad \text{et} \quad \{P_i, Q_j\}_{\vec{x}} = -\delta_{ij},$$

c'est-à-dire la transformation préserve les parenthèses de Poisson fondamentales.

DÉMONSTRATION. Montrons que (1) implique (2). Nous avons vu qu'une transformation est canonique si et seulement si sa matrice Jacobienne, $A_{ij} = \frac{\partial x_i}{\partial X_j}$, est symplectique pour tout t :

$$(A)^T \mathcal{J} A = \mathcal{J};$$

or les gradients se transforment selon la loi :

$$\nabla_{\vec{X}} F = (A)^T \nabla_{\vec{x}} f,$$

en effet (notation indices répétés implique indices additionnés)

$$\frac{\partial F}{\partial X_i}(\vec{X}, t) = \frac{\partial}{\partial X_i} f(\vec{x}(\vec{X}, t), t) = \frac{\partial f}{\partial x_j}(\vec{x}, t) \frac{\partial x_j}{\partial X_i} = \frac{\partial f}{\partial x_j} A_{ji}.$$

Mais alors par définition de parenthèse de Poisson :

$$\begin{aligned} \{F, G\}_{\vec{X}} &= (\nabla_{\vec{X}} F)^T \mathcal{J} \nabla_{\vec{X}} G = ((A)^T \nabla_{\vec{x}} f)^T \mathcal{J} (A)^T \nabla_{\vec{x}} g = (\nabla_{\vec{x}} f)^T A \mathcal{J} (A)^T \nabla_{\vec{x}} g \\ &= (\nabla_{\vec{x}} f)^T \mathcal{J} \nabla_{\vec{x}} g = \{f, g\}_{\vec{x}}. \end{aligned}$$

L'implication (2) \Rightarrow (3) est triviale, car (III.8.7) est un cas particulier de (III.8.6) et pour les variables (\vec{p}, \vec{q}) les relations (III.8.7) ont déjà été vérifiées (cfr. remarque 8.2).

Il nous reste à prouver que (3) implique (1), cela est obtenu en montrant que les relations (III.8.7) impliquent que la matrice Jacobienne est symplectique. Nous avons vu que une matrice est symplectique si et seulement si sa transposée est symplectique, donc nous pouvons nous limiter à prouver que $A\mathcal{J}(A)^T = \mathcal{J}$; il s'agit par un simple calcul⁹ de montrer que pour, une transformation quelconque, la matrice $A\mathcal{J}(A)^T$ a la structure à blocs suivante :

$$A\mathcal{J}(A)^T = \begin{pmatrix} \{P_i, P_j\} & \{P_i, Q_j\} \\ \{Q_i, P_j\} & \{Q_i, Q_j\} \end{pmatrix},$$

mais alors les relations (III.8.7) nous permettent de conclure que :

$$A\mathcal{J}(A)^T = \begin{pmatrix} \mathbb{O} & -\delta_{ij} \\ \delta_{ij} & \mathbb{O} \end{pmatrix} = \mathcal{J}.$$

Ce qui termine la démonstration. \square

9. Invariants intégraux

Dans ce paragraphe nous donnons une dernière caractérisation des transformations canoniques à l'aide des formes différentielles, pour simplifier le résultat nous allons nous placer dans \mathbb{R}^{2m} . Si nous observons que la 2-forme

$$\omega = \sum_{j=1}^m dp_j \wedge dq_j,$$

nous donne la somme des projections sur les plans (p_j, q_j) du volume dans l'espace de phases et si nous nous rappelons du Théorème de Liouville (cfr. Théorème II.4.2), alors le résultat teollo suivant n'est pas étonnant

⁹Soit $A_{ij} = \partial X_i / \partial x_j$ la matrice Jacobienne de la transformation de coordonnées et écrivons-la sous la forme $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, avec a, b, c et d quatre matrices quadratiques $m \times m$. Plus précisément $a_{ij} = \partial P_i / \partial p_j$, $b_{ij} = \partial P_i / \partial q_j$, $c_{ij} = \partial Q_i / \partial p_j$ et $d_{ij} = \partial Q_i / \partial q_j$. Mais alors

$$A\mathcal{J}(A)^T = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \mathbb{O} & -\mathbb{I} \\ \mathbb{I} & \mathbb{O} \end{pmatrix} \begin{pmatrix} (a)^T & (c)^T \\ (b)^T & (d)^T \end{pmatrix} = \begin{pmatrix} b & -a \\ d & -c \end{pmatrix} \begin{pmatrix} (a)^T & (c)^T \\ (b)^T & (d)^T \end{pmatrix} = \begin{pmatrix} b(a)^T - a(b)^T & b(c)^T - a(d)^T \\ d(a)^T - c(b)^T & d(c)^T - c(d)^T \end{pmatrix}.$$

On obtient le résultat en écrivant en composantes chaque bloc $m \times m$, par exemple

$$(b(c)^T - a(d)^T)_{ij} = \sum_l \left[\frac{\partial P_i}{\partial q_l} \frac{\partial Q_j}{\partial p_l} - \frac{\partial P_i}{\partial p_l} \frac{\partial Q_j}{\partial q_l} \right] = \{P_i, Q_j\}.$$

THÉOREME 9.1. *Une transformation de coordonnées $(\vec{p}, \vec{q}) = (\vec{p}(\vec{P}, \vec{Q}), \vec{q}(\vec{P}, \vec{Q}))$ est complètement canonique si et seulement si :*

$$(III.9.1) \quad \sum_{j=1}^m dp_j \wedge dq_j = \sum_{j=1}^m dP_j \wedge dQ_j .$$

DÉMONSTRATION. La démonstration de la condition suffisante suit immédiatement de la condition de Lie :

$$\sum_{j=1}^m (p_j dq_j - P_j dQ_j) = df ,$$

puisque la transformation est indépendante du temps nous avons $\tilde{d} = d$, mais alors $d^2 f = 0$ et la condition (III.9.1) suit automatiquement.

Vice versa, puisque (III.9.1) est équivalente à :

$$d \left(\sum_{j=1}^m (p_j dq_j - P_j dQ_j) \right) = 0 ,$$

c'est-à-dire la 1-forme $p_j dq_j - P_j dQ_j$ est fermée, alors nous pouvons conclure grâce au Lemme de Poincaré (dans \mathbb{R}^k toute forme fermée est exacte), qu'il existe une fonction f telle que :

$$\sum_{j=1}^m (p_j dq_j - P_j dQ_j) = df ,$$

et encore la condition de Lie nous permet de conclure. □

Transformations proches de l'identité : Série de Lie.

Ce chapitre est en phase de construction.

1. Introduction

Dans le cours d'analyse vous avez vu la notion de dérivée d'une fonction d'une variable : son signe indique les directions selon lesquelles la fonction croît ou décroît. En passant à des fonctions de plusieurs variables, la généralisation naturelle est de considérer les dérivées partielles par rapport aux variables, en réalité cette notion n'est pas la plus générale. En effet si \vec{v} est un vecteur de \mathbb{R}^n nous pouvons mesurer le taux de croissance (ou décroissance) de la fonction selon la direction donnée par le vecteur, c'est la *dérivée directionnelle* :

$$(IV.1.1) \quad \frac{\partial f}{\partial \vec{v}}(\vec{x}) := \lim_{t \rightarrow 0} \frac{f(\vec{x} + t\vec{v}) - f(\vec{x})}{t|\vec{v}|},$$

c'est-à-dire on calcule la limite de l'incrément infinitésimal le long de la direction donnée par \vec{v} .

Cette idée peut être généralisée pour une direction qui change avec le (i.e. dépend du) point \vec{x} , soit donc $\vec{v}(\vec{x})$ un champ de vecteurs, on appelle *dérivée de Lie* associée à \vec{v} l'opérateur :

$$(IV.1.2) \quad L_{\vec{v}(\vec{x})} := \sum_{i=1}^n v_i \frac{\partial}{\partial x_i}.$$

Par définition la dérivée de Lie est *linéaire* et satisfait à la *règle de Leibniz* :

$$L_{\vec{v}(\vec{x})}(fg) = fL_{\vec{v}(\vec{x})}g + gL_{\vec{v}(\vec{x})}f.$$

Si le champ de vecteurs est associé à une équation différentielle alors la dérivée de Lie peut être calculée en partant de la notion de flot de l'équation différentielle. Soit donc

$$(IV.1.3) \quad \dot{\vec{x}} = \vec{v}(\vec{x}),$$

alors nous avons

PROPOSITION 1.1. *La dérivée de Lie (selon la direction de \vec{v}) d'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est donnée par :*

$$(L_{\vec{v}(\vec{x})}f)(\vec{x}) = \frac{d}{dt}f \circ \Phi^t(\vec{x})|_{t=0},$$

où Φ^t est le flot au temps t de (IV.1.3) et $f \circ g$ représente la composition de f et g .

DÉMONSTRATION. Un simple calcul nous donne :

$$\frac{d}{dt}f \circ \Phi^t(\vec{x})|_{t=0} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \frac{d\Phi_i^t}{dt}(\vec{x})|_{t=0},$$

mais par définition de flot : $\frac{d\Phi_i^t}{dt}(\vec{x})|_{t=0} = v_i(\Phi^t(\vec{x})|_{t=0} = v_i(\vec{x})$, car $\Phi^t(\vec{x})|_{t=0} = \vec{x}$.
Donc :

$$\frac{d}{dt}f \circ \Phi^t(\vec{x})|_{t=0} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} v_i(\vec{x}) = (L_{\vec{v}(\vec{x})}f)(\vec{x}).$$

□

REMARQUE 1.2. Cette proposition permet de donner une autre caractérisation des intégrales premières ; une fonction f est une intégrale première si et seulement si sa dérivée de Lie vaut zéro.

REMARQUE 1.3. Si le champ de vecteurs est hamiltonien, c'est-à-dire il existe une fonction \mathcal{H} telle que $\vec{v} = \mathcal{J}\nabla\mathcal{H}$, alors la dérivée de Lie d'une fonction f est liée à la parenthèse de Poisson :

$$(L_{\vec{v}(\vec{x})}f)(\vec{x}) = \frac{d}{dt}f \circ \Phi^t(\vec{x})|_{t=0} = \sum_{i=1}^{2m} \frac{\partial f}{\partial x_i} (\mathcal{J}\nabla\mathcal{H})_i(\vec{x}) = (\nabla f)^T \mathcal{J}\nabla\mathcal{H} = \{f, \mathcal{H}\}.$$

Soit maintenant deux champs de vecteurs \vec{v}_1 et \vec{v}_2 , avec les flots associés, Φ_1^t et Φ_2^t . En général suivre d'abord le flot Φ_1^t pour un temps t et après Φ_2^s pour un temps s , donnera un résultat différent de faire le parcours en ordre inversé, autrement dit les flots *ne commutent pas*.

EXEMPLE 1.4. Considérons un premier flot $\Phi_1^t(\vec{x}) = (x_1 \cos t - x_2 \sin t, x_1 \sin t + x_2 \cos t)$, si nous introduisons la matrice $R(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$ alors il est immédiat que ce flot correspond à une rotation d'un angle t :

$$\Phi_1^t(\vec{x}) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Et un deuxième flot $\Phi_2^t(\vec{x}) = (x_1 + t, x_2)$, i.e. une translation de t le long de la première composante.

Avec la dérivée de Lie nous pouvons mesurer le *dégré de non-commutation* des deux flots. Pour cela prenons une quelconque fonction régulière $f : \mathbb{R}^n \rightarrow \mathbb{R}$, alors la non-commutativité est mesurée par :

$$(IV.1.4) \quad (\Delta f)(t, s, \vec{x}) = f \circ \Phi_2^s \Phi_1^t(\vec{x}) - f \circ \Phi_1^t \Phi_2^s(\vec{x}).$$

Clairement nous avons

$$(\Delta f)(0, 0, \vec{x}) = 0,$$

car les flots ne déplacent rien au temps zéro. Nous allons chercher le développement de Taylor de cette fonction de deux variables. Par définition :

$$\frac{d}{dt}g \circ \Phi_j^t(\vec{x})|_{t=0} = (L_{\vec{v}_j(\vec{x})}g)(\vec{x}),$$

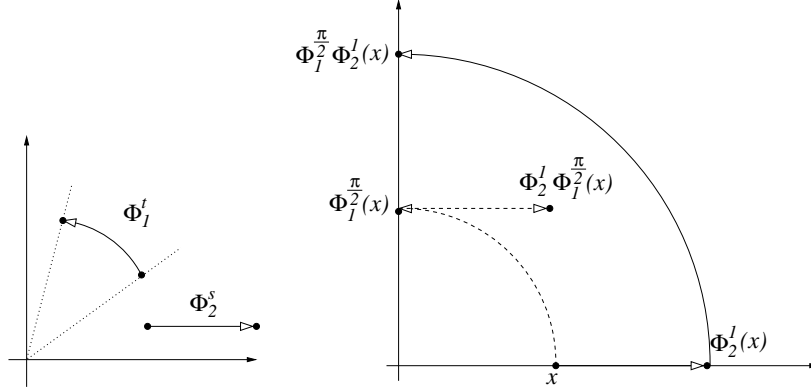


FIG. 1. Un exemple de deux flots qui ne commutent pas :
 $\Phi_1^{\frac{\pi}{2}} \Phi_2^1(\vec{x}) \neq \Phi_2^1(\vec{x}) \Phi_1^{\frac{\pi}{2}}$.

et donc pour les dérivées premières :

$$\begin{aligned} \frac{\partial}{\partial t}(\Delta f)(t, s, \vec{x})|_{t=0, s=0} &= \frac{d}{dt}(f \circ \Phi_2^s)(\Phi_1^t(\vec{x}))|_{t=0, s=0} - \frac{d}{dt}f \circ \Phi_1^t(\Phi_2^s(\vec{x}))|_{t=0, s=0} \\ &= (L_{\vec{v}_1}(\vec{x})(f \circ \Phi_2^s))(\vec{x})|_{s=0} - (L_{\vec{v}_1}(\Phi_2^s(\vec{x}))f)(\Phi_2^s(\vec{x}))|_{s=0} \\ &= (L_{\vec{v}_1}(\vec{x})f)(\vec{x}) - (L_{\vec{v}_1}(\vec{x})f)(\vec{x}) = 0, \end{aligned}$$

et de façon similaire pour $\frac{\partial}{\partial s}(\Delta f)(t, s, \vec{x})|_{t=0, s=0} = 0$. Donc au premier ordre la série de Taylor est nulle ; le calcul du deuxième ordre est plus long, nous affirmons que le seul terme non nul est :

$$\frac{\partial^2}{\partial t \partial s} \Delta f(t, s, \vec{x})|_{t=0, s=0}.$$

Un petit calcul permet de montrer que :

$$\frac{\partial^2}{\partial t \partial s} \Delta f(t, s, \vec{x})|_{t=0, s=0} = (L_{[\vec{v}_1, \vec{v}_2]}f)(\vec{x}),$$

où $[\vec{v}_1, \vec{v}_2]$ est le *commutateur* des champs \vec{v}_1 et \vec{v}_2 , défini par :

$$(IV.1.5) \quad [\vec{v}_1, \vec{v}_2]_i := \sum_{l=1}^n \left(\vec{v}_{1,l} \frac{\partial \vec{v}_{2,i}}{\partial x_l} - \vec{v}_{2,l} \frac{\partial \vec{v}_{1,i}}{\partial x_l} \right) = L_{\vec{v}_1}(\vec{v}_2)_i - L_{\vec{v}_2}(\vec{v}_1)_i,$$

et nous avons la propriété remarquable que :

$$(IV.1.6) \quad L_{[\vec{v}_1, \vec{v}_2]} = [L_{\vec{v}_1}, L_{\vec{v}_2}].$$

En partant du résultat suivant (voir Arnol'd, méthodes mathématiques de la mécanique classique) : deux flots commutent si et seulement si leur commutateur vaut zéro, et des remarques précédentes nous pouvons conclure que les flots associés aux deux hamiltoniens \mathcal{H}_1 et \mathcal{H}_2 commutent si et seulement si $\{\mathcal{H}_1, \mathcal{H}_2\} = 0$. Deux fonctions telles que $\{f, g\} = 0$ sont dites *en involution*.

Théorie d'Hamilton–Jacobi et variables Action–Angle 1–dim.

1. Introduction et exemples

Nous avons déjà vu que le flot hamiltonien définit une transformation canonique (cfr. Théorème III.4.9), nous pouvons aussi démontrer que le nouvel hamiltonien, \mathcal{K} , est identiquement zéro. Nous allons ici nous occuper du problème inverse : chercher un changement (canonique) de coordonnées tel que dans les nouvelles coordonnées le nouvel hamiltonien soit identiquement zéro, de cette façon l'intégration du mouvement sera triviale, et le changement de coordonnées nous permettra d'exprimer les variables initiales en fonction du temps (intégration des équations d'Hamilton) :

$$\begin{aligned}\mathcal{H}, \vec{p}, \vec{q} &\Rightarrow \mathcal{K} \equiv 0, \vec{P}(t) = \vec{\eta}, \vec{Q}(t) = \vec{\xi} \\ &\Rightarrow \vec{p}(t) = \vec{p}(\vec{\eta}, \vec{\xi}, t), \vec{q}(t) = \vec{q}(\vec{\eta}, \vec{\xi}, t).\end{aligned}$$

Supposons de ne pas être dans un voisinage d'un point singulier de \mathcal{H} , $\nabla_{\vec{x}}\mathcal{H} \neq 0$, puisque la transformation qui échange \vec{p} avec \vec{q} (à un signe près) est canonique nous pouvons supposer que : $\nabla_{\vec{p}}\mathcal{H} \neq 0$. Nous pouvons chercher cette transformation avec une fonction génératrice de la forme $S = S(\vec{q}, \vec{\eta}, t)$ (vieilles variables \vec{q} et nouvelles variables \vec{P}). Mais alors nous avons :

$$(V.1.1) \quad p_i = \frac{\partial S}{\partial q_i}$$

$$(V.1.2) \quad \xi_i = \frac{\partial S}{\partial \eta_i}$$

$$(V.1.3) \quad \mathcal{K} = \mathcal{H} + \frac{\partial S}{\partial t},$$

c'est-à-dire avec les contraintes choisies ($\mathcal{K} \equiv 0$) :

$$(V.1.4) \quad \mathcal{H}(\nabla_{\vec{q}}S, \vec{q}, t) + \frac{\partial S}{\partial t} = 0,$$

pour tout $\vec{\eta}$ dans un ouvert convenable de \mathbb{R}^m , et la condition de non-dégénérescence :

$$(V.1.5) \quad \det \left(\frac{\partial^2 S}{\partial q_i \partial \eta_j} \right) \neq 0.$$

L'équation (V.1.4) est appelée *équation d'Hamilton–Jacobi*, il s'agit d'une équation aux dérivées partielles du premier ordre non-linéaire, les variables indépendantes sont les q_1, \dots, q_m et le temps t . Sa solution va au delà des méthodes vues dans le cours Équations Différentielles Ordinaires, nous nous limiterons ici à la recherche d'une solution qui dépend de $m+1$ constantes arbitraires, autant que le nombre des variables, dite *intégrale complète*. Une solution de la forme $S(\vec{q}, \vec{\eta}, t) + \eta_0$, avec η_0

et $\vec{\eta}$ constantes arbitraires, est exactement une intégrale complète, elle est appelée aussi *fonction principale de Hamilton*.

Nous pouvons donc démontrer les résultats suivants :

THÉOREME 1.1. *Soit donné l'hamiltonien $\mathcal{H}(\vec{p}, \vec{q}, t)$ et soit $S(\vec{q}, \vec{\eta}, t)$ une intégrale complète de l'équation d'Hamilton–Jacobi (V.1.4), dépendante de $m+1$ constantes arbitraires et qui vérifie la condition (V.1.5). Alors les solutions des équations d'Hamilton pour \mathcal{H} sont obtenues de :*

$$(V.1.6) \quad p_j = \frac{\partial S}{\partial q_j}, \quad \xi_j = \frac{\partial S}{\partial \eta_j} \quad \forall j = 1, \dots, m,$$

où ξ_1, \dots, ξ_m sont des constantes arbitraires.

DÉMONSTRATION. La fonction S vérifie la condition de non-dégénérescence et donc le système de nouvelles coordonnées $(\vec{\xi}, \vec{\eta})$ est canonique, l'équation d'Hamilton–Jacobi, implique que le nouvel hamiltonien vaut identiquement zéro et donc les équations d'Hamilton deviennent :

$$\begin{cases} \dot{\xi}_j &= -\frac{\partial \mathcal{K}}{\partial \xi_j} \equiv 0 \\ \dot{\eta}_j &= \frac{\partial \mathcal{K}}{\partial \eta_j} \equiv 0. \end{cases}$$

Donc les variables $\vec{\xi}$ et $\vec{\eta}$ sont des constantes et le résultat est démontré. \square

REMARQUE 1.2. *Si l'hamiltonien ne dépend pas du temps de façon explicite alors la solution S peut être cherchée de la forme $((\vec{p}, \vec{q}) \mapsto (\vec{\eta}, \vec{\xi}))$:*

$$(V.1.7) \quad S = -E(\vec{\eta})t + W(\vec{q}, \vec{\eta}),$$

où $\vec{\eta}$ est un vecteur (m -dim) constant et $E(\vec{\eta})$ est une fonction (au moins \mathcal{C}^2) telle que $\nabla_{\vec{\eta}} E \neq 0$ (en général nous pouvons choisir $\eta_1 = E$). La fonction W , dite fonction caractéristique de Hamilton, vérifie la relation (aussi dite équation d'Hamilton–Jacobi) :

$$(V.1.8) \quad \mathcal{H}(\nabla_{\vec{q}} W, \vec{q}) = E(\vec{\eta}),$$

qui nous permet d'identifier E avec l'énergie du système. La variable $\vec{\xi}$ peut être obtenue via la fonction génératrice S ,

$$\vec{\xi} = \frac{\partial S}{\partial \vec{\eta}} = -\frac{\partial E}{\partial \vec{\eta}} t + \frac{\partial W}{\partial \vec{\eta}}.$$

Puisque le nouvel hamiltonien vérifie par construction $\mathcal{K} = 0$ nous avons $\vec{\xi}(t) = \text{const}$, qui substitué dans l'équation précédente permet de calculer la dépendance de \vec{q} de la variable temps, une fois que nous avons obtenu et inversé la fonction W (voir exemple 1.4).

Cela peut résulter très difficile mais il existe une façon alternative pour obtenir le même résultat. Considérons le changement de coordonnées $(\vec{p}, \vec{q}) \mapsto (\vec{\alpha}, \vec{\beta})$, avec $\vec{\alpha} = \vec{\eta}$, engendré par la fonction $W(\vec{q}, \vec{\alpha})$. Nous pouvons observer que :

$$0 \neq \det \left(\frac{\partial^2 S}{\partial q_i \partial \eta_j} \right) = \det \left(\frac{\partial^2 W}{\partial q_i \partial \eta_j} \right) = \det \left(\frac{\partial^2 W}{\partial q_i \partial \alpha_j} \right),$$

et donc le nouveau système de variables $(\vec{\alpha}, \vec{\beta})$ est canonique et nous pouvons penser la fonction W comme une transformation canonique de coordonnées, dont le nouvel

hamiltonien est $\tilde{\mathcal{H}} = E(\vec{\alpha})$ (voir (V.1.8)) et donc les équations d'Hamilton sont :

$$(V.1.9) \quad \begin{cases} \dot{\alpha}_j &= -\frac{\partial \tilde{\mathcal{H}}}{\partial \beta_j} = 0 \\ \dot{\beta}_j &= \frac{\partial \tilde{\mathcal{H}}}{\partial \alpha_j} = \gamma_j(\vec{\alpha}), \end{cases}$$

qui peuvent être explicitement résolues :

$$(V.1.10) \quad \alpha_j(t) = \alpha_j(0) \quad \text{et} \quad \beta_j(t) = \gamma(\vec{\alpha}(0))_j t + \beta_j(0).$$

Les variables $(\vec{\alpha}, \vec{\beta})$ sont donc liées aux variables $(\vec{\eta}, \vec{\xi})$ par : $\vec{\alpha} = \vec{\eta}$ et $\vec{\beta} = \vec{\xi} + t \partial_{\vec{\eta}} E$.

REMARQUE 1.3. De façon similaire, si \mathcal{H} ne dépend pas explicitement du temps et des variables q_1, \dots, q_k , $k < m$, alors nous pouvons chercher la fonction caractéristique de Hamilton sous la forme :

$$W = \sum_{j=1}^k \alpha_j q_j + W_0(q_{k+1}, \dots, q_m, \alpha_1, \dots, \alpha_m).$$

EXEMPLE 1.4 (Oscillateur harmonique). Considérons un oscillateur harmonique 1-dimensionnel :

$$\mathcal{H}_{oh}(p, q) = \frac{1}{2m} p^2 + \frac{m\omega^2}{2} q^2,$$

l'équation d'Hamilton–Jacobi a la forme :

$$\frac{1}{2m} \left[\left(\frac{\partial S}{\partial q} \right)^2 + m^2 \omega^2 q^2 \right] + \frac{\partial S}{\partial t} = 0.$$

Nous allons introduire

$$S(q, E, t) = W(q, E) - Et,$$

(où la variable E joue le rôle de η ou α dans le schéma précédent) et l'équation pour W devient :

$$\frac{1}{2m} \left[\left(\frac{\partial W}{\partial q} \right)^2 + m^2 \omega^2 q^2 \right] = E,$$

d'où :

$$W(q, E) = \sqrt{2mE} \int_{q_0}^q \sqrt{1 - \frac{m\omega^2 x^2}{2E}} dx.$$

Si nous posons $q_0 = 0$ alors l'intégrale est explicite :

$$W(q, E) = \sqrt{\frac{mE}{2}} \left[q \sqrt{1 - \frac{m\omega^2 q^2}{2E}} + \sqrt{\frac{2E}{m\omega^2}} \arcsin \left(q \sqrt{\frac{m\omega^2}{2E}} \right) \right],$$

d'où (ξ étant la variable conjuguée à E) :

$$\xi = \frac{\partial S}{\partial E} = -t + \frac{1}{\omega} \arcsin \left(q \sqrt{\frac{m\omega^2}{2E}} \right),$$

qui permet d'obtenir

$$q = \sqrt{\frac{2E}{m\omega^2}} \sin(\omega t + \omega \xi),$$

et ensuite p par la définition $p = \partial_q W$.

La deuxième façon de procéder nous aurait donné :

$$\beta = \frac{\partial W}{\partial E} = \frac{1}{\omega} \arcsin \left(q \sqrt{\frac{m\omega^2}{2E}} \right),$$

et l'équation d'hamilton pour β suivante

$$\dot{\beta} = \frac{\partial E}{\partial \alpha} = 1,$$

donc $\beta(t) = \beta(0) + t$ et finalement :

$$\begin{cases} p &= \frac{\partial W}{\partial q} = \sqrt{2mE} \cos(\omega\beta) \\ q &= \sqrt{\frac{2E}{m\omega^2}} \sin(\omega\beta). \end{cases}$$

EXEMPLE 1.5 (Systèmes 1-dimensionnels conservatifs). *Considérons un point de masse m soumis à une force conservative à laquelle on associe une énergie potentielle $V(x)$, l'hamiltonien du système est donc :*

$$\mathcal{H}(p, x) = \frac{p^2}{2m} + V(x).$$

L'équation d'Hamilton-Jacobi est (pour la fonction caractéristique) :

$$\frac{1}{2m} \left(\frac{\partial W}{\partial x} \right)^2 + V(x) = E,$$

qui peut être intégrée (nous avons donc identifié α_1 avec l'énergie E) :

$$W(x, E) = \sqrt{2m} \int_{x_0}^x \sqrt{E - V(\xi)} d\xi.$$

La transformation canonique engendrée est :

$$\begin{cases} p &= \frac{\partial W}{\partial x} = \sqrt{2m[E - V(x)]} \\ \beta &= \frac{\partial W}{\partial E} = \sqrt{\frac{m}{2}} \int_{x_0}^x \frac{d\xi}{\sqrt{E - V(\xi)}}. \end{cases}$$

2. Séparation des variables

La *séparation des variables* est une technique qui permet de résoudre explicitement dans beaucoup de cas l'équation d'Hamilton-Jacobi ; cette technique est applicable quand, par exemple, l'hamiltonien peut être écrit comme une somme de m hamiltoniens chacun dépendant seulement d'un couple de variables (p_j, q_j) :

$$(V.2.1) \quad \mathcal{H}(\vec{p}, \vec{q}) = \mathcal{H}_1(p_1, q_1) + \cdots + \mathcal{H}_m(p_m, q_m).$$

Alors l'équation d'Hamilton-Jacobi admet la solution suivante :

$$(V.2.2) \quad W = \sum_{j=1}^m W_j(p_j, q_j),$$

où chaque fonction W_j est solution de l'équation

$$(V.2.3) \quad \mathcal{H}_j \left(\frac{\partial W_j}{\partial q_j}, q_j \right) = e_j(\alpha_j),$$

avec e_j fonction régulière arbitraire, mais telle que :

$$(V.2.4) \quad E(\alpha_1, \dots, \alpha_m) = \sum_{j=1}^m e_j(\alpha_j),$$

car de $S = W - Et$ et de l'équation d'Hamilton–Jacobi complète nous obtenons :

$$E = \mathcal{H}_1 \left(\frac{\partial S}{\partial q_1}, q_1 \right) + \dots + \mathcal{H}_m \left(\frac{\partial S}{\partial q_m}, q_m \right) = \mathcal{H}_1 \left(\frac{\partial W_1}{\partial q_1}, q_1 \right) + \dots + \mathcal{H}_m \left(\frac{\partial W_m}{\partial q_m}, q_m \right).$$

Un simple exemple est donné par un oscillateur harmonique dans \mathbb{R}^3 , l'hamiltonien du système est :

$$\mathcal{H} = \frac{p_1^2 + p_2^2 + p_3^2}{2m} + \frac{m}{2} (\omega_1^2 + \omega_2^2 + \omega_3^2),$$

c'est-à-dire

$$\mathcal{H} = \mathcal{H}_{oh}(p_1, q_1) + \mathcal{H}_{oh}(p_2, q_2) + \mathcal{H}_{oh}(p_3, q_3),$$

où $\mathcal{H}_{oh}(p_j, q_j)$ est l'hamiltonien de l'oscillateur harmonique 1-dim vu dans l'exemple 1.4.

Plus généralement nous pouvons étudier le cas où l'hamiltonien dépend de la variable q_1 et $\frac{\partial S}{\partial q_1}$ à travers une fonction $h_1(\frac{\partial S}{\partial q_1}, q_1)$, indépendante des autres variables et du temps, c'est-à-dire l'hamiltonien est de la forme :

$$(V.2.5) \quad \mathcal{H} = \mathcal{H}(h_1(p_1, q_1), p_2, \dots, p_m, q_2, \dots, q_m, t).$$

Mais alors l'équation d'Hamilton–Jacobi devient :

$$(V.2.6) \quad \mathcal{H} \left(h_1 \left(\frac{\partial S}{\partial q_1}, q_1 \right), \frac{\partial S}{\partial q_2}, \dots, \frac{\partial S}{\partial q_m}, q_2, \dots, q_m, t \right) + \frac{\partial S}{\partial t} = 0,$$

et nous pouvons en chercher une solution de la forme :

$$(V.2.7) \quad S = S_1(q_1, \alpha_1) + S'(q_2, \dots, q_m, \alpha_2, \dots, \alpha_m, t),$$

car :

$$(V.2.8) \quad h_1 \left(\frac{\partial S}{\partial q_1}, q_1 \right) = e_1(\alpha_1) \\ \mathcal{H} \left(e_1(\alpha_1), \frac{\partial S'}{\partial q_2}, \dots, \frac{\partial S'}{\partial q_m}, q_2, \dots, q_m, t \right) + \frac{\partial S'}{\partial t} = 0.$$

La première équation est une EDO du premier ordre d'où nous pouvons déduire S_1 par intégration; la deuxième équation est encore une équation aux dérivées partielles mais qui dépend de m variables au lieu de $m+1$: nous avons baissé l'ordre. Si nous pouvons itérer la méthode $m+1$ fois nous aurons déterminé une intégrale complète de l'équation d'Hamilton–Jacobi *par quadrature*, c'est-à-dire avec des intégrations, dans ce cas le système hamiltonien est dit *séparable*.

Cette technique est très puissante mais il s'avère que souvent l'hamiltonien n'est pas dans la forme demandée, néanmoins parfois on peut se ramener à cette forme avec un simple changement de coordonnées.

EXEMPLE 2.1. *Considérons l'exemple d'un système d'un point de masse m qui se déplace dans l'espace 3-dimensionnel soumis à une force centrale conservative. Si la position du point est déterminée par le vecteur $\vec{x} = (x, y, z)$ alors le lagrangien est :*

$$\mathcal{L}(\vec{x}, \dot{\vec{x}}) = \frac{m}{2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - V(r),$$

où $r = \sqrt{x^2 + y^2 + z^2}$ et V est l'énergie potentielle associée à la force centrale conservative. L'hamiltonien correspondant est :

$$\mathcal{H}(\vec{p}, \vec{x}) = \frac{\vec{p}^2}{2m} + V(r),$$

qui n'est pas dans la forme à variables séparables; nous nous proposons de montrer que en passant en coordonnées sphériques, l'hamiltonien devient à variables séparables.

Les coordonnées sphériques en dimension 3 permettent de passer des coordonnées rectangulaires (x, y, z) aux coordonnées (r, θ, ϕ) grâce aux formules suivantes (voir Fig. 1) :

$$(V.2.9) \quad \begin{cases} x &= r \cos \phi \cos \theta \\ y &= r \cos \phi \sin \theta \\ z &= r \sin \phi. \end{cases}$$

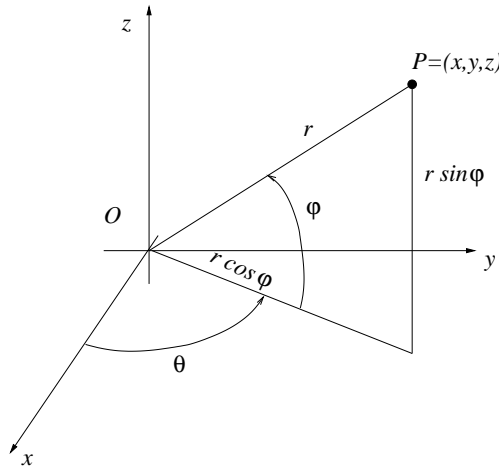


FIG. 1. Coordonnées sphériques en \mathbb{R}^3 .

Le calcul de la vitesse dans ces coordonnées est obtenu en dérivant par rapport au temps (V.2.9) :

$$(V.2.10) \quad \begin{cases} \dot{x} &= \dot{r} \cos \phi \cos \theta - r \dot{\phi} \sin \phi \cos \theta - r \dot{\theta} \cos \phi \sin \theta \\ \dot{y} &= \dot{r} \cos \phi \sin \theta - r \dot{\phi} \sin \phi \sin \theta + r \dot{\theta} \cos \phi \cos \theta \\ \dot{z} &= \dot{r} \sin \phi + r \dot{\phi} \cos \phi, \end{cases}$$

et pour le carré de la vitesse :

$$\dot{\vec{x}}^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2 = \dot{r}^2 + r^2 \dot{\phi}^2 + r^2 \dot{\theta}^2 \cos^2 \phi,$$

ce qui permet de récrire le lagrangien comme :

$$\mathcal{L}(r, \phi, \theta, \dot{r}, \dot{\phi}, \dot{\theta}) = \frac{m}{2} \left(\dot{r}^2 + r^2 \dot{\phi}^2 + r^2 \dot{\theta}^2 \cos^2 \phi \right) - V(r).$$

Nous voulons maintenant introduire l'hamiltonien en coordonnées sphériques, et pour cela il faut utiliser la définition (i.e. transformée de Legendre du lagrangien).

Par définition nous avons :

$$\begin{aligned} p_r &:= \frac{\partial \mathcal{L}}{\partial \dot{r}} = m\dot{r} \\ p_\phi &:= \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = mr^2 \dot{\phi} \\ p_\theta &:= \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = mr^2 \cos^2 \phi \dot{\theta}, \end{aligned}$$

et donc :

$$\begin{aligned} \mathcal{H}(p_r, p_\phi, p_\theta, r, \phi, \theta) &= \vec{p} \cdot \dot{\vec{x}} - \mathcal{L} = p_r \dot{r} + p_\phi \dot{\phi} + p_\theta \dot{\theta} \\ &= \frac{p_r^2}{m} + \frac{p_\phi^2}{mr^2} + \frac{p_\theta^2}{mr^2 \cos^2 \phi} - \mathcal{L} \\ &= \frac{p_r^2}{2m} + \frac{p_\phi^2}{2mr^2} + \frac{p_\theta^2}{2mr^2 \cos^2 \phi} + V(r). \end{aligned}$$

Il est facile maintenant de montrer que l'hamiltonien est à variables séparables ; l'équation d'Hamilton–Jacobi est :

$$\frac{1}{2m} \left[\left(\frac{\partial S}{\partial r} \right)^2 + \frac{1}{r^2} \left(\frac{\partial S}{\partial \phi} \right)^2 + \frac{1}{r^2 \cos^2 \phi} \left(\frac{\partial S}{\partial \theta} \right)^2 \right] + \frac{\partial S}{\partial t} + V(r) = 0,$$

or une intégrale complète est donnée par :

$$S = W_1(r) + W_2(\phi) + W_3(\theta) - Et,$$

où les fonctions W_j vérifient l'équation :

$$\frac{1}{2m} \left(\frac{dW_1}{dr} \right)^2 + \frac{1}{2mr^2} \left(\frac{dW_2}{d\phi} \right)^2 + \frac{1}{2mr^2 \cos^2 \phi} \left(\frac{dW_3}{d\theta} \right)^2 - E + V(r) = 0.$$

Or cette équation peut être “divisée en deux parties” :

$$\frac{1}{2m} \left(\frac{dW_1}{dr} \right)^2 - E + V(r) + \frac{c}{r^2} = 0 \quad \text{et} \quad \frac{1}{2m} \left(\frac{dW_2}{d\phi} \right)^2 + \frac{1}{2m \cos^2 \phi} \left(\frac{dW_3}{d\theta} \right)^2 = c,$$

dont la première ne dépend que de la variable r et la deuxième des variables ϕ et θ ; mais cette deuxième équation peut encore être simplifiée :

$$\left(\frac{dW_3}{d\theta} \right)^2 = 2m \cos^2 \phi \left[c - \frac{1}{2m} \left(\frac{dW_2}{d\phi} \right)^2 \right] \Rightarrow \left(\frac{dW_3}{d\theta} \right)^2 = c' \quad \text{et} \quad \left(\frac{dW_2}{d\phi} \right)^2 = 2mc - \frac{c'}{\cos^2 \phi}.$$

Chaque équation pour W_j dépend d'une seule variable et donc en principe résoluble avec une intégration : l'hamiltonien est donc à variables séparables.

3. Variables Angle–Action en dimension 1

Considérons un système hamiltonien avec un degré de liberté autonome, $\mathcal{H}(p, q)$. Nous savons que la valeur numérique de l'hamiltonien est une constante du mouvement qui correspond à l'énergie mécanique du système : $\mathcal{H}(p, q) = E$, cela est équivalent à dire que les trajectoires sont les courbes de niveau, $\gamma_E = \{(p, q) : \mathcal{H}(p, q) = E\}$, pour toute valeur de l'énergie. Supposons maintenant que pour des valeurs de l'énergie dans un intervalle ouvert toutes ces courbes γ_E dans l'espace des phases soient *simples* (i.e. sans auto intersections), *fermées*, *connexes* (autrement on pourra se limiter à ne considérer qu'une seule composante connexe à la fois) et *non-singulières* (c'est-à-dire le gradient de l'hamiltonien évalué sur n'importe quelle courbe, ne s'annule pas : $\nabla \mathcal{H}|_{\gamma_E} \neq 0$) (voir Fig. 2). Puisque les courbes sont

fermées le mouvement est périodique : après un certain temps, T_E , qui en général dépend de la valeur de l'énergie, les variables $(p(t), q(t))$ reviennent à leurs valeurs initiales ; ces orbites sont donc appelées : *orbites de libration* ou *d'oscillation*.

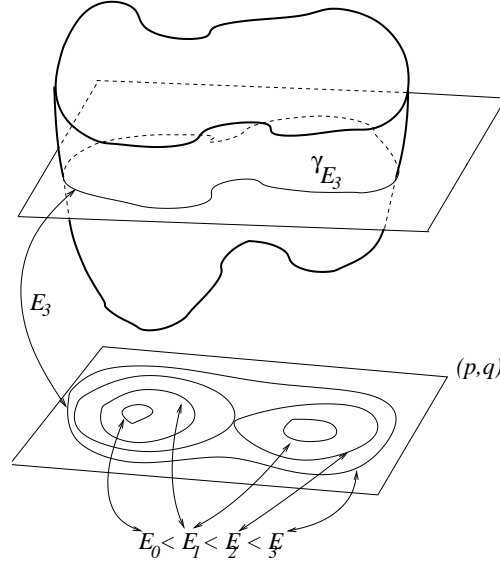


FIG. 2. Courbes d'énergie constante dans \mathbb{R}^2 .

Dans ces hypothèses chaque courbe γ_E est difféomorphe à un cercle qui a la même aire que celle bornée par la courbe γ_E . Puisque chaque courbe γ_E est rectifiable, si λ_E est sa longueur on peut introduire un paramètre s qui décrit la courbe, $\gamma_E = (p_E(s), q_E(s))$, alors le cercle sera décrit par un angle $\psi = 2\pi \frac{s}{\lambda_E}$ et chaque point sera repéré par $p = R_E \cos \psi$ et $q = R_E \sin \psi$, où le rayon R_E est choisi de façon à avoir un cercle de même aire. Cela nous permet finalement d'introduire des nouvelles coordonnées, de (p, q) à (R_E, ψ) dans lesquelles la courbe γ_E a une équation très simple $R_E = \text{const}$; en général ces nouvelles coordonnées ne sont pas canoniques comme nous le montrons dans l'exemple 3.2.

Il est donc naturel se demander s'il existe des coordonnées canoniques avec des propriétés similaires à celles vues pour ψ et R_E , notamment :

- (1) une variable est un *angle*, c'est-à-dire elle augmente de 2π chaque fois que la courbe γ_E est parcourue ;
- (2) l'autre variable dépend seulement de l'énergie, c'est-à-dire elle identifie univoquement chaque courbe γ_E ;
- (3) grâce à la propriété précédente le nouvel hamiltonien exprimé dans les nouvelles variables dépendra seulement de la deuxième variable appelée *action*.

DÉFINITION 3.1. *S'il existe une transformation complètement canonique :*

$$(V.3.1) \quad \begin{cases} p &= p(A, \phi) \\ q &= q(A, \phi), \end{cases}$$

où la dépendance de ϕ est 2π -périodique, qui vérifie :

$$(V.3.2) \quad E = \mathcal{H}(p(A, \phi), q(A, \phi)) = \mathcal{K}(A) \quad \text{et} \quad \int_{\gamma_E} d\phi = 2\pi,$$

alors le système sera dit complètement canoniquement intégrable et les nouvelles variables Action–Angle.

Si nous arrivons à mettre un système en variables action–angle alors nous pourrions l’intégrer directement, car :

$$(V.3.3) \quad \frac{dA}{dt} = -\frac{\partial \mathcal{K}}{\partial \phi} = 0,$$

donc l’action est une constante du mouvement, et

$$(V.3.4) \quad \frac{d\phi}{dt} = \frac{\partial \mathcal{K}}{\partial A} := \omega(A),$$

l’angle évolue de façon linéaire :

$$(V.3.5) \quad \phi(t) = \phi(0) + \omega(A(0))t.$$

Remarquons qu’il existe une autre classe de systèmes hamiltoniens de dimension 1 qui admettent des variables action–angle. Supposons en effet que l’hamiltonien dépende périodiquement de la variable q : $\mathcal{H}(p, q + \lambda) = \mathcal{H}(p, q)$ pour tout (p, q) . Supposons de plus que les courbes d’énergie constante γ_E soient *simples* et *non-singulières* quand E varie dans un intervalle ouvert. Si ces courbes sont aussi fermées alors le mouvement est encore une oscillation, par contre si elles ne sont pas fermées alors le mouvement sera une *rotation*, dans ce cas ces courbes se réduisent au graphe d’une fonction régulière et périodique $p = p(q, E)$. Nous supposons aussi que les courbes varient en fonction de l’énergie E , $\frac{\partial p}{\partial E} \neq 0$ (voir Fig. 3)

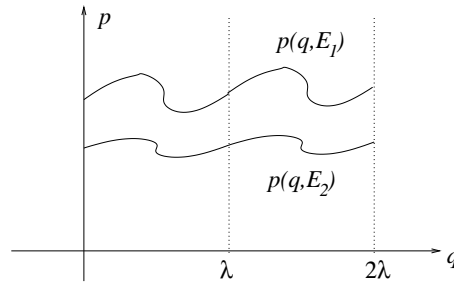


FIG. 3. Courbes d’énergie constante dans \mathbb{R}^2 pour un hamiltonien périodique.

Les variables action–angle, A, ϕ , pour ces systèmes devront satisfaire aux propriétés suivantes :

$$(V.3.6) \quad p = p(A, \phi) \Rightarrow p(A, \phi) = p(A, \phi + 2\pi) \quad \forall (A, \phi)$$

$$(V.3.7) \quad q = q(A, \phi) \Rightarrow q(A, \phi) = q(A, \phi + 2\pi) + \lambda \quad \forall (A, \phi).$$

Un système hamiltonien peut “contenir” les deux classes de mouvement, un exemple classique est le pendule : si la vitesse initiale est suffisamment grande, alors le pendule fera des rotations, si par contre la vitesse est petite alors le mouvement sera des oscillations. Nous verrons la construction des variables action–angle pour

le pendule dans l'exemple 3.5, maintenant nous nous penchons sur un exemple plus simple : l'oscillateur harmonique.

EXEMPLE 3.2 (Variables action-angle pour l'oscillateur harmonique). *Considérons l'oscillateur harmonique, $\mathcal{H}(p, q) = \frac{p^2}{2m} + \frac{m\omega^2 q^2}{2}$, alors les courbes γ_E sont les ellipses $\frac{p^2}{2mE} + \frac{m\omega^2 q^2}{2E} = 1$, dont l'aire vaut $\frac{2\pi}{\omega}E$.*

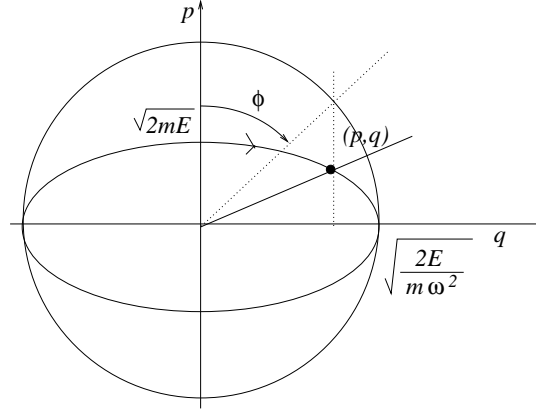


FIG. 4. Courbe d'énergie constante pour l'oscillateur harmonique.

Nous cherchons un cercle de rayon R_E tel que : $\pi R_E^2 = \frac{2\pi}{\omega}E$, et un angle comme montré dans la Fig. 4 :

$$p = \sqrt{2mE} \cos \phi \quad \text{et} \quad q = \sqrt{\frac{2E}{m\omega^2}} \sin \phi,$$

car il est facile de vérifier que (p, q) décrivent l'ellipse quand ϕ varie entre 0 et 2π ; le changement de coordonnées est donc donné par :

$$p = \sqrt{m\omega} R_E \cos \phi \quad \text{et} \quad q = \sqrt{\frac{1}{m\omega}} R_E \sin \phi.$$

Or ces nouvelles coordonnées ne sont pas canoniques, pour montrer cela il est suffisant de calculer la parenthèse de Poisson $\{\phi, R_E\}_{(p,q)}$ et montrer qu'elle ne vaut pas 1. Pour éviter de devoir exprimer ϕ et R_E en fonction de p et q , nous pouvons calculer la parenthèse de Poisson $\{q, p\}_{(\phi, R_E)}$ et montrer qu'elle est différente de 1, et puisque nous avons vu que les transformations canoniques forment un groupe, nous pouvons conclure qu'une transformation est canonique si et seulement si son inverse l'est.

$$\begin{aligned} \{q, p\}_{(\phi, R_E)} &= \frac{\partial q}{\partial \phi} \frac{\partial p}{\partial R_E} - \frac{\partial p}{\partial \phi} \frac{\partial q}{\partial R_E} \\ &= \frac{1}{\sqrt{m\omega}} R_E \cos \phi \sqrt{m\omega} \cos \phi + \sqrt{m\omega} R_E \sin \phi \frac{1}{\sqrt{m\omega}} \sin \phi \\ &= R_E \neq 1. \end{aligned}$$

Or pour obtenir une transformation canonique nous devons modifier les relations trouvées, en particulier puisque nous voulons que le nouvel hamiltonien soit

une fonction de la seule variable d'action, nous pouvons définir l'action comme une fonction arbitraire de R_E , $A = f(R_E)$. D'après les calculs de la parenthèse de Poisson il s'avère un choix approprié :

$$R_E = \sqrt{2A}.$$

De cette façon le changement de coordonnées devient :

$$p = \sqrt{2m\omega A} \cos \phi \quad \text{et} \quad q = \sqrt{\frac{2A}{m\omega}} \sin \phi,$$

et puisque l'énergie est préservée, le nouvel hamiltonien vérifie $\mathcal{K}(A) = E$, c'est-à-dire :

$$\mathcal{K}(A) = E = \frac{\omega}{2} R_E^2 = \omega A.$$

Il nous reste à prouver que la transformation est canonique :

$$\begin{aligned} \{q, p\}_{(\phi, A)} &= \frac{\partial q}{\partial \phi} \frac{\partial p}{\partial A} - \frac{\partial p}{\partial \phi} \frac{\partial q}{\partial A} \\ &= \sqrt{\frac{2A}{m\omega}} \cos \phi \frac{1}{2} \sqrt{\frac{2m\omega}{A}} \cos \phi + \sqrt{2m\omega A} \sin \phi \frac{1}{2} \sqrt{\frac{2}{m\omega A}} \sin \phi \\ &= 1. \end{aligned}$$

Nous avons donc trouvé un changement de coordonnées des variables (p, q) à des nouvelles variables (A, ϕ) tel que le nouvel hamiltonien dépend seulement de la variable A , qui géométriquement correspond à l'aire de l'ellipse divisée par 2π :

$$A = \frac{E}{\omega} = \left(\frac{2\pi E}{\omega} \right) \frac{1}{2\pi}.$$

Dans l'exemple 1.4 nous avons déjà déterminé une transformation de coordonnées qui permettait de transformer l'hamiltonien de l'oscillateur harmonique dans les variables (p, q) à des nouvelles variables (E, β) telles que le nouvel hamiltonien dépendait seulement de E . Cette transformation n'est rien d'autre que la transformation à variables action-angle, où $E = \omega A$ et $\beta = \phi$, donc :

$$F(q, A) = A \arcsin \left(\sqrt{\frac{\omega m}{2A}} q \right) + q \sqrt{\frac{m\omega A}{2}} \sqrt{1 - \frac{m\omega}{2A} q^2},$$

est la fonction génératrice (du deuxième type) qui réalise la transformation canonique.

Nous allons maintenant introduire la méthode qui permet de construire en général une transformation canonique à variables action-angle pour les mouvements de rotation et d'oscillation. Pour faire cela nous allons utiliser la méthode des fonctions génératrices; nous cherchons donc une fonction $F(q, A)$ (vieilles variables q et nouvelles variables P , mais que nous dénotons par A comme action), telle que :

$$(V.3.8) \quad p = \frac{\partial F}{\partial q} \quad \text{et} \quad \phi = \frac{\partial F}{\partial A},$$

et qui vérifie la condition de non-dégénérescence :

$$(V.3.9) \quad \det \frac{\partial^2 F}{\partial q \partial A} \neq 0.$$

Nous avons vu que dans le cas des rotations et des oscillations, nous pouvons exprimer localement la variable p comme un fonction de (q, E) , $p = \hat{p}(q, E)$, puisque la variable action que l'on cherche devra satisfaire à la relation : $E = \mathcal{K}(A)$, c'est-à-dire le nouvel hamiltonien dépendra seulement de la variable action et de plus l'énergie est préservée, alors nous pouvons utiliser comme fonction génératrice (quitte à devoir encore définir la fonction $\mathcal{K}(A) = E$ et à démontrer que $\frac{d\mathcal{K}}{dA} \neq 0$, pour pouvoir inverser la relation entre énergie et action) la fonction suivante :

$$(V.3.10) \quad F(q, A) := \int_{q_0}^q \hat{p}(x, \mathcal{K}(A)) dx,$$

qui correspond à l'intégrale de la 1-forme $p dq$ le long la courbe γ_E . En effet en dérivant cette fonction par rapport à q nous trouvons :

$$\frac{\partial F}{\partial q} = \hat{p}(q, \mathcal{K}(A)),$$

et puisque localement sur γ_E nous avons $\mathcal{K}(A) = E$ et $p = \hat{p}(q, E)$, nous avons bien obtenu que :

$$\frac{\partial F}{\partial q} = p.$$

En outre

$$\det \frac{\partial^2 F}{\partial q \partial A} = \det \frac{\partial \hat{p}}{\partial E} \frac{d\mathcal{K}}{dA} \neq 0.$$

Finalement si nous introduisons la variation de la fonction F quand la variable q parcourt un tour le long la courbe γ_E :

$$\Delta F(A) := \int_{\gamma_E} p(q, A) dq,$$

où $p(q, A) = \hat{p}(q, \mathcal{K}(A))$ et $\mathcal{K}(A) = E$, alors nous obtenons de la deuxième équation (V.3.8) :

$$\int_{q_0}^q d\phi|_{A=const} = \frac{\partial}{\partial A} \int_{q_0}^q p(x, A) dx,$$

et finalement

$$\int_{\gamma_E} d\phi = \frac{d}{dA} \Delta F(A).$$

Observons que la fonction génératrice est *multivoque*, c'est-à-dire à plusieurs valeurs, cela découle du fait que la 1-forme $p dq$ n'est pas exacte. Géométriquement la signification de $\Delta F(A)$ est l'aire bornée par la courbe γ_E dans le cas des oscillations, ou l'aire délimitée par le graphe de γ_E et l'axe des q pour le cas des rotations. Pour satisfaire la dernière condition (V.3.2) il faudra imposer :

$$\frac{d}{dA} \Delta F(A) = 2\pi.$$

Nous pouvons donc définir

DÉFINITION 3.3 (Variable Action). *On appelle variable action la quantité :*

$$(V.3.11) \quad A := \frac{1}{2\pi} \int_{\gamma_E} p dq = \frac{\mathcal{A}(E)}{2\pi},$$

où nous avons dénoté par $\mathcal{A}(E)$ l'aire délimitée par la courbe γ_E .

Nous pouvons maintenant vérifier que : $\frac{d\mathcal{K}}{dA} \neq 0$. Pour faire cela nous calculons la dérivée $\frac{dA}{dE}$, géométriquement cela correspond à dire que dans la Fig. 2 quand on varie le niveau d'énergie alors l'aire de la courbe délimitée doit aussi changer. Plus précisément nous pouvons démontrer que (voir Fig. 5) si nous augmentons l'énergie d'un petit incrément δE alors la différence d'aire bornée par les courbes $\mathcal{H} = E$ et $\mathcal{H} = E + \delta E$ dans un voisinage d'un point $x = (p, q)$ est donnée par :

$$\mathcal{A}_x(E + \delta E) - \mathcal{A}_x(E) = ds dh,$$

où ds est l'élément infinitésimal sur la courbe γ_E autour de x et dh est le déplacement infinitésimal entre γ_E et $\gamma_{E+\delta E}$. Si on observe que le gradient de \mathcal{H} est perpendiculaire à la courbe γ_E (c'est-à-dire tous les points de γ_E se déplacent dans la direction donnée par ce vecteur), alors nous avons que $\delta E = |\nabla_x \mathcal{H}| dh$. Pour calculer la différence entre les aires il faut additionner toutes ces contributions locales :

$$\mathcal{A}(E + \delta E) - \mathcal{A}(E) = \iint_{\Gamma_E} ds dh = \int_{\gamma_E} ds \frac{\delta E}{|\nabla_x \mathcal{H}|},$$

où Γ_E est "l'anneau" délimité par les courbes γ_E et $\gamma_{E+\delta E}$, et nous avons éliminé l'intégrale sur le déplacement dh car il dépend du point x . Finalement car par hypothèse sur γ_E nous avons $\nabla_x \mathcal{H} \neq 0$ nous pouvons conclure que :

$$\frac{\mathcal{A}(E + \delta E) - \mathcal{A}(E)}{\delta E} = \int_{\gamma_E} \frac{ds}{|\nabla_x \mathcal{H}|} + \mathcal{O}(\delta E)^2,$$

et donc en passant à la limite $\delta E \rightarrow 0$ nous trouvons :

$$\frac{d\mathcal{A}}{dE} \neq 0,$$

et donc grâce à la définition d'action :

$$\frac{dA}{dE} = \frac{1}{2\pi} \frac{d\mathcal{A}}{dE} \neq 0$$

ce qui implique que $\frac{\partial \mathcal{K}}{\partial A} \neq 0$, car $\mathcal{K} = E$.

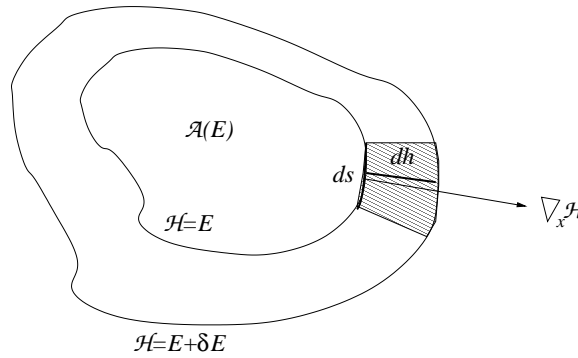


FIG. 5. Géométrie pour la démonstration de $\frac{dA}{dE} \neq 0$.

Nous avons donc démontré le résultat suivant

THÉOREME 3.4 (Variables action-angle en dimension 1). *Tout système hamiltonien à un degré de liberté et autonome, $\mathcal{H}(p, q)$, avec mouvement de rotation et/ou oscillation, admet des variables action-angle et donc il est complètement canoniquement intégrable.*

Nous concluons ce Chapitre avec un dernier exemple.

EXEMPLE 3.5 (Variables action-angle pour le pendule). *Considérons un pendule (point de masse m , suspendu à un fil de longueur l dans un champ de gravité g), dont l'hamiltonien est :*

$$\mathcal{H}(p, q) = \frac{p^2}{2ml^2} - mgl \cos q,$$

si nous posons $e = \frac{E}{mgl}$ (rapport entre l'énergie mécanique et l'énergie potentielle due à la gravitation) alors pour $|e| < 1$ le pendule a des oscillations.

La variable d'action est obtenue en intégrant la fonction $p(q, e)$ sur un niveau d'énergie donné γ_E :

$$A = \frac{4}{2\pi} \int_0^{q_M} \sqrt{2m^2 l^3 g} \sqrt{e + \cos q} dq,$$

où nous avons dénoté par q_M le point sur la courbe γ_E tel que $\cos q_M + e = 0$, il s'agit d'un point d'inversion, c'est-à-dire un point où le mouvement change de direction car la vitesse (proportionnelle à p) s'annule et ensuite change de signe. Le facteur 4 revient à découper la courbe γ_E en quatre parties symétriques, dont la contribution est identique à l'intégrale complète.

Cette intégrale ne peut pas être calculée en termes de fonctions élémentaires, nous pouvons l'exprimer en termes de fonctions elliptiques :

$$A = ml\sqrt{gl} \frac{8}{\pi} [(k^2 - 1)\mathbb{K}(k) + \mathbb{E}(k)],$$

avec $k^2 = \frac{e+1}{2}$ et $\mathbb{K}(k)$ et $\mathbb{E}(k)$ sont les fonctions elliptiques de premier et deuxième type.

Si $|e| > 1$ le pendule a des mouvements de rotation, l'action a une forme similaire mais cette fois $q_M = \pi$, en réalité il n'y a pas de points d'inversion car la vitesse (p) est toujours du même signe. Encore une fois on peut l'exprimer seulement en termes de fonctions elliptiques.

CHAPITRE VI

Intégration par quadratures : Liouville, Arnol'd et Action–Angle.

1. Introduction

Dans ce Chapitre nous allons présenter un des sujets qui ont déterminé le grand développement et application de la théorie d'Hamilton. En fait pour intégrer un système de $2l$ équations différentielles ordinaires du premier ordre il faut connaître en général $2l$ intégrales premières.

EXEMPLE 1.1. *Si on considère le système de 2 équations :*

$$\begin{cases} \dot{x} &= f(x, y) \\ \dot{y} &= g(x, y) \end{cases}$$

alors une façon de procéder serait de déterminer deux fonctions, I_1 et I_2 , telles que :

$$\left. \frac{d}{dt} I_j \right|_{flow} = 0 \quad \text{pour } j = 1, 2;$$

dans le cas affirmatif nous pourrions fixer la valeur des deux intégrales premières : $I_1(x, y) = c_1$ et $I_2(x, y) = c_2$, ensuite déterminer un lien entre $x(t)$ et $y(t)$ en utilisant la première relation :

$$I_1(x, y) = c_1 \quad \Rightarrow \quad x(t) = F_1(y(t), c_1),$$

et finalement substituer dans la deuxième relation pour déterminer $y(t)$

$$I_2(x, y) = c_2 \quad \Rightarrow \quad I_2(F_1(y(t), c_1), y) = c_2 \quad \Rightarrow \quad y(t) = F_2(c_1, c_2).$$

Après $x(t)$ pourra être déterminé en utilisant à nouveau la première relation.

Cependant si le système est hamiltonien le fait que la structure symplectique est préservée par le flot hamiltonien entraîne une conséquence remarquable : il est suffisant de connaître seulement l intégrales premières (indépendantes et en involution, c'est-à-dire que les crochets de Poisson sont nuls $\{f, g\} = 0$.) pour résoudre les équations d'Hamilton.

Ce résultat contenu dans le Théorème de Liouville permettra, sous certaines hypothèses supplémentaires, de démontrer l'existence de variables *action–angle* pour un système hamiltonien en dimension quelconque; puisque les actions sont des intégrales premières nous pouvons en général résoudre les équations d'Hamilton. Cela n'implique pas que le mouvement résultant sera “simple”, comme il sera montré dans des cas particuliers dans le paragraphe 5.

2. Intégration par quadrature : le Théorème de Liouville

Le but de cette section est d'introduire et démontrer le Théorème de Liouville. Avant de faire cela nous allons montrer sur un exemple les conséquences de ce Théorème, considérons donc un pendule dont l'hamiltonien en coordonnées canoniques (et normalisées afin d'éliminer tous les paramètres) est :

$$(VI.2.1) \quad \mathcal{H}(p, q) = \frac{p^2}{2} - \cos q;$$

puisque le système est autonome, $\frac{\partial \mathcal{H}}{\partial t} = 0$, l'hamiltonien est une constante du mouvement, autrement dit il est une intégrale première. Considérons donc les ensembles de niveau de $\mathcal{H} = \text{const}$, c'est-à-dire les lieux de points (p, q) tels que (voir Fig. 1) :

$$(VI.2.2) \quad \frac{p^2}{2} - \cos q = E.$$

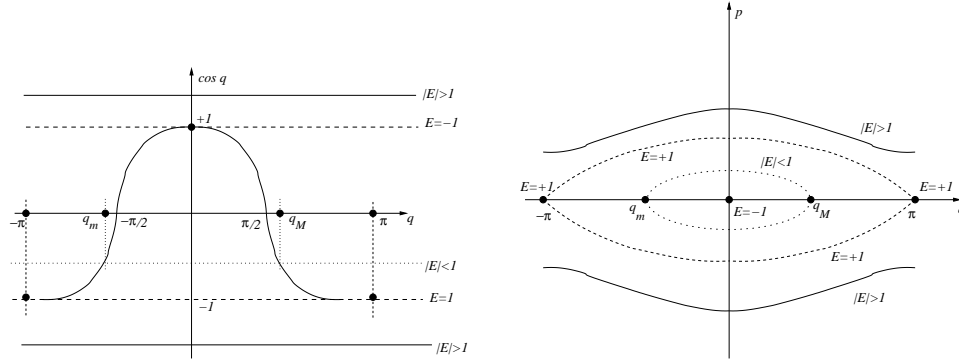


FIG. 1. Ensembles de niveau pour le pendule. À gauche la courbe $y = \cos q$ avec les courbes $y = -E$, à droite les “courbes” $p^2 - 2 \cos q = 2E$.

Si la valeur de la constante E est, en valeur absolue, plus grande que un, $|E| > 1$, alors la fonction \cos ne rencontre jamais la droite $y = -E$, autrement dit, pour aucun point $q \in (-\pi, \pi]$ nous avons $\cos q = -E$, et plus précisément $\cos q < -E$ (si $-E > 1$) ou bien $-E < \cos q$ (si $-E < -1$). Donc nous pouvons extraire la racine et écrire ces courbes de niveau comme :

$$(VI.2.3) \quad p_{-}(q) = -\sqrt{2(E + \cos q)} \quad \text{et} \quad p_{+}(q) = \sqrt{2(E + \cos q)},$$

avec les fonctions $p_{\pm}(q)$ régulières (elles sont au moins \mathcal{C}^1 et la dérivée ne s'annule jamais). Ces courbes sont périodiques de période 2π et elles correspondent à des *rotations* : la variable p (vitesse) ne s'annule jamais et la variable angulaire augmente toujours (si on ne la regarde pas mod 2π) : le mouvement est donc sur une courbe qui est *difféomorphe* à la droite réelle \mathbb{R} .

Si $|E| < 1$, alors il existe deux points $q_m < 0 < q_M$ tels que : $\cos q_m = \cos q_M = -E$, nous pouvons toujours exprimer l'ensemble de niveau avec les courbes $p_{\pm}(q)$ sauf que cette fois ces courbes ne sont pas régulières : aux points q_m et q_M la dérivée première n'est pas définie. En réalité nous pouvons toujours décrire ces ensembles de niveau avec des courbes régulières mais qui ne seront pas des graphes. Ces courbes correspondent à des *oscillations*, la variable p (vitesse) oscille entre deux valeurs

$p_M > 0 > p_m$, données par $p_M = \sqrt{2(E+1)}$ et $p_m = -\sqrt{2(E+1)}$, et la variable q entre $q_m < 0 < q_M$. Donc le mouvement est périodique et il est sur une courbe difféomorphe à un cercle (ou un *tore*) \mathbb{T} de dimension 1.

Considérons maintenant les cas limites : $E = \pm 1$. Si $E \rightarrow -1^+$ alors les points q_m et q_M tendent vers le point $q = 0$ et les points p_m et p_M vers le point $p = 0$; pour $E = -1$ la seule intersection entre $y = \cos q$ et $y = -E$ est le point $q = 0$ et donc l'ensemble de niveau $\mathcal{H} = -1$ est réduit à un seul point $(p, q) = (0, 0)$. On peut vérifier que ce point est un point d'équilibre pour les équations d'Hamilton. Si $E \rightarrow +1^-$, alors les points q_m et q_M tendent respectivement vers les points $-\pi$ et π ; en ces points les courbes p_{\pm} sont continues mais pas différentiable, il est facile de vérifier que :

$$p'_+(-\pi) = p'_-(\pi) = 1 \quad \text{et} \quad p'_+(\pi) = p'_-(-\pi) = -1.$$

Ces orbites ne sont ni des rotations ni des oscillations, elles sont appelées *séparatrices*, car elles séparent les deux types de mouvements réguliers.

Dans cet exemple du pendule nous avons vu que la dynamique régulière d'un système hamiltonien de dimension 1 est difféomorphe à la droite réelle, mouvements non bornés, ou bien à un 1-tore, mouvements bornés, de plus *une seule* intégrale première \mathcal{H} a été suffisante pour intégrer le mouvement. Le Théorème de Liouville nous dit que ce comportement est tout à fait général pour un système hamiltonien de dimension quelconque : la dynamique est toujours sur une variété obtenue comme produit de droites et de tores. Puisque le théorème est valable en dimension quelconque on aura besoin d'un nombre plus grand d'intégrales premières et elles devront satisfaire certaines conditions de compatibilité.

THÉOREME 2.1 (Liouville). *Soit donc donné un système hamiltonien avec l degrés de liberté autonome, $\mathcal{H}(\vec{p}, \vec{q})$, et supposons qu'il existe l intégrales premières, $f_1(\vec{p}, \vec{q}), \dots, f_l(\vec{p}, \vec{q})$ indépendantes et en involution. Considérons l'ensemble de niveau des intégrales premières :*

$$(VI.2.4) \quad M_{\vec{a}} := \{\vec{x} = (\vec{p}, \vec{q}) \in \mathbb{R}^{2l} : f_j(\vec{p}, \vec{q}) = a_j, j = 1, \dots, l\},$$

avec $\vec{a} = (a_1, \dots, a_l) \in \mathbb{R}^l$ est un vecteur fixé.

Alors si $M_{\vec{a}}$ n'est pas vide nous avons :

- (1) $M_{\vec{a}}$ est une sous-variété régulière de dimension l de \mathbb{R}^{2l} , elle est invariante pour le flot hamiltonien $\Phi_{\mathcal{H}}$ et pour les flots Φ_{f_j} , $j = 1, \dots, l$ associés aux "hamiltoniens" f_j , $j = 1, \dots, l$;
- (2) Les flots Φ_{f_j} , $j = 1, \dots, l$ commutent entre eux;
- (3) Si en plus

$$(VI.2.5) \quad \det \left(\frac{\partial f_i}{\partial p_j} \right) \neq 0,$$

alors il existe (localement) une fonction $S = S(\vec{f}, \vec{q}, t)$ telle que :

$$(VI.2.6) \quad \left(\sum_{i=1}^l p_i dq_i - \mathcal{H}(\vec{p}, \vec{q}) dt \right) \Big|_{(\vec{p}, \vec{q}) \in M_{\vec{a}}} = dS(\vec{a}, \vec{q}, t),$$

et qui est une intégrale complète de l'équation d'Hamilton-Jacobi pour \mathcal{H} . Le système est donc intégrable par quadrature.

Avant de passer à la démonstration remarquons quelques points importants.

REMARQUE 2.2. La condition que les l intégrales premières, $f_1(\vec{p}, \vec{q}), \dots, f_l(\vec{p}, \vec{q})$ soient indépendantes signifie que les gradients $\nabla_{(p,q)} f_1(\vec{p}, \vec{q}), \dots, \nabla_{(p,q)} f_l(\vec{p}, \vec{q})$ soient l vecteurs linéairement indépendants. Comme nous avons déjà vu la condition que deux fonctions soient en involution implique que, vues comme fonctions d'Hamilton, leur flots commutent.

REMARQUE 2.3. Puisque le système est autonome la fonction d'Hamilton est une intégrale première qui peut donc être comptée parmi les l intégrales premières f_j . De toute façon grâce à la condition d'involution, \mathcal{H} et toutes les f_j sont constantes sur leur flot mais aussi sur les flots engendrés par les autres fonctions.

Puisque $\mathcal{H}|_{M_{\vec{a}}}$ est constant, la relation (VI.2.6) devient (localement) :

$$(VI.2.7) \quad \sum_{j=1}^l p_j dq_j|_{M_{\vec{a}}} = dW(\vec{a}, \vec{q}),$$

avec

$$(VI.2.8) \quad dW(\vec{a}, \vec{q}) = dS(\vec{a}, \vec{q}, t) + E(\vec{a})t,$$

et $E(\vec{a})$ est la valeur de \mathcal{H} sur $M_{\vec{a}}$.

REMARQUE 2.4. La condition (VI.2.5) n'est pas restrictive, car la condition d'indépendance linéaire des fonctions f_j , assure l'existence de l coordonnées canoniques, x_1, \dots, x_l , telles que :

$$(VI.2.9) \quad \det \left(\frac{\partial(f_1, \dots, f_l)}{\partial(x_1, \dots, x_l)} \right) \neq 0,$$

mais quitte à échanger certaines variables q_j avec les relatives variables p_j (transformation canonique) nous pouvons supposer que (VI.2.9) soit satisfaite avec les variables \vec{p} .

DÉMONSTRATION. Les propriétés (1) et (2) suivent facilement de l'hypothèse d'indépendance linéaire des gradients : $M_{\vec{a}}$ est donc une sous-variété régulière de dimension l , car l'espace tangent est orthogonal à l'espace formé par les gradients qui a dimension l dans \mathbb{R}^{2l} . Grâce à l'hypothèse d'involution les flots commutent et donc la variété est invariante par les flots.

La condition (VI.2.5) assure (théorème de la fonction implicite) que localement il existe l fonctions régulières $\hat{p}_1(\vec{f}, \vec{q}), \dots, \hat{p}_l(\vec{f}, \vec{q})$ telles que (nous avons dénoté par $\vec{f} = (f_1, \dots, f_l)$) :

$$(VI.2.10) \quad p_j = \hat{p}_j(\vec{f}, \vec{q}) \quad \text{et} \quad f_j = (\vec{p}, \vec{q}) = a_j \quad \forall j = 1, \dots, l.$$

Or la relation (VI.2.6) est équivalente (voir remarque 2.3) à :

$$(VI.2.11) \quad \sum_{j=1}^l \hat{p}_j(\vec{f}, \vec{q}) dq_j = dW(\vec{f}, \vec{q}),$$

et l'existence d'une telle fonction W est assurée (c'est-à-dire la forme $\hat{p}_j(\vec{f}, \vec{q}) dq_j$ est exacte et localement cela correspond à dire que la forme est fermée) si pour tout $j, k = 1, \dots, l$:

$$(VI.2.12) \quad \frac{\partial \hat{p}_k}{\partial q_j} = \frac{\partial \hat{p}_j}{\partial q_k},$$

autrement dit si la matrice $B_{kj} = \frac{\partial \hat{p}_k}{\partial q_j}$ est symétrique ; ce que nous allons démontrer maintenant.

En dérivant par rapport à q_k la deuxième relation de (VI.2.10) nous obtenons (notation indices répétés égal indices additionnés) :

$$(VI.2.13) \quad 0 = \frac{\partial f_i}{\partial p_j} \frac{\partial \hat{p}_j}{\partial q_k} + \frac{\partial f_i}{\partial q_k},$$

pour simplifier nous allons introduire les matrices :

$$A_{ij} = \frac{\partial f_j}{\partial p_j} \quad \text{et} \quad C_{ij} = \frac{\partial f_i}{\partial q_j},$$

ce qui permet de récrire (VI.2.13) comme :

$$(VI.2.14) \quad 0 = A_{ij} B_{jk} + C_{ik} \Rightarrow 0 = AB + C.$$

Nous devons donc démontrer que la matrice $B = -A^{-1}C$ est symétrique. Un simple calcul nous donne (notez que : $-(C)^T = (AB)^T = (B)^T(A)^T$, donc $(B)^T = -(C)^T((A)^T)^{-1}$) :

$$(VI.2.15) \quad \begin{aligned} B - (B)^T &= -A^{-1}C + (C)^T((A)^T)^{-1} \\ &= -A^{-1}(C(A)^T - A(C)^T)((A)^T)^{-1}. \end{aligned}$$

Donc la matrice B est symétrique si et seulement si $C(A)^T - A(C)^T = 0$, or si nous écrivons en composantes cette matrice nous trouvons (toujours notation sur les indices répétés) :

$$(VI.2.16) \quad (C(A)^T - A(C)^T)_{ij} = C_{il}A_{jl} - A_{il}C_{jl} = \frac{\partial f_i}{\partial q_l} \frac{\partial f_j}{\partial p_l} - \frac{\partial f_i}{\partial p_l} \frac{\partial f_j}{\partial q_l} = \{f_i, f_j\} = 0,$$

grâce à l'hypothèse d'involution. Donc la matrice B est symétrique et localement l'existence de la fonction W est prouvée, sur $M_{\vec{a}}$ nous avons donc :

$$(VI.2.17) \quad W(\vec{f}, \vec{q}) = \int_{\vec{q}_0}^{\vec{q}} \sum_{j=1}^l \hat{p}_j(\vec{f}, \vec{x}) d\xi_j,$$

où le point \vec{q}_0 est un point arbitraire de $M_{\vec{a}}$ et l'intégration est faite sur un chemin arbitraire qui relie \vec{q}_0 à \vec{q} (toujours appartenant à $M_{\vec{a}}$). Nous avons pu passer à des valeurs de \vec{f} non constantes puisque la valeur de \vec{a} est arbitraire.

Considérons maintenant le nouvel hamiltonien dans les nouvelles variables :

$$(VI.2.18) \quad \hat{\mathcal{H}}(\vec{f}, \vec{q}) := \mathcal{H}(\vec{\hat{p}}(\vec{f}, \vec{q}), \vec{q}),$$

si on fixe $\vec{f} = \vec{a}$ alors de la condition de symétrie de B on trouve (notation sur les indices) :

$$(VI.2.19) \quad \begin{aligned} \frac{\partial \hat{\mathcal{H}}}{\partial q_i} &= \frac{\partial \mathcal{H}}{\partial q_i} + \frac{\partial \mathcal{H}}{\partial p_j} \frac{\partial \hat{p}_j}{\partial q_i} = \frac{\partial \mathcal{H}}{\partial q_i} + \frac{\partial \mathcal{H}}{\partial p_j} \frac{\partial \hat{p}_i}{\partial q_j} \\ &= -\dot{p}_i + \frac{\partial \hat{p}_i}{\partial q_j} \dot{q}_j = -\dot{p}_i + \frac{d}{dt} \hat{p}_i = 0, \end{aligned}$$

où la dernière égalité a été obtenue car par définition $\hat{p}_i = p_i$ et $\vec{f} = \text{const.}$ Donc le nouvel hamiltonien *ne dépend pas* des variables \vec{q} et donc il dépend seulement des

variables \vec{f} :

$$(VI.2.20) \quad \mathcal{H}(\vec{p}(\vec{f}, \vec{q}), \vec{q}) = \hat{\mathcal{H}}(\vec{f}).$$

Finalement si on pose

$$(VI.2.21) \quad S(\vec{f}, \vec{q}, t) := W(\vec{f}, \vec{q}) - \hat{\mathcal{H}}(\vec{f})t,$$

alors (VI.2.6) est trivialement satisfaite et S est solution de l'équation d'Hamilton–Jacobi. En fait (par définition de W) :

$$p_i = \frac{\partial W}{\partial q_i} = \frac{\partial S}{\partial q_i},$$

et

$$\det \left(\frac{\partial^2 W}{\partial q_i \partial f_j} \right) = \det \left(\frac{\partial p_i}{\partial f_j} \right) \neq 0,$$

et finalement (VI.2.21) implique que :

$$\mathcal{H} + \frac{\partial S}{\partial t} = 0.$$

Donc S est une intégrale complète puisqu'elle dépend des l constantes arbitraires a_1, \dots, a_l . \square

Le Théorème de Liouville implique que les intégrales premières f_1, \dots, f_l peuvent jouer le rôle de nouvelles coordonnées canoniques avec les variables $\vec{\beta}$ définies par :

$$(VI.2.22) \quad \beta_j := \frac{\partial W}{\partial f_j} \quad j = 1, \dots, l.$$

La fonction W est donc la fonction génératrice d'une transformation complètement canonique des variables (\vec{p}, \vec{q}) aux variables $(\vec{f}, \vec{\beta})$. Le nouvel hamiltonien est fonction des seules variables \vec{f} et donc les équations d'Hamilton sont :

$$(VI.2.23) \quad \begin{cases} \dot{\vec{f}} = 0 \\ \dot{\vec{\beta}} = \nabla_{\vec{f}} \hat{\mathcal{H}}(\vec{f}), \end{cases}$$

d'où nous obtenons le fait que \vec{f} sont des constantes et

$$(VI.2.24) \quad \vec{\beta}(t) = \vec{\beta}(0) + \nabla_{\vec{f}} \hat{\mathcal{H}}(\vec{f}(0))t.$$

3. Le Théorème d'Arnol'd

Le Théorème de Liouville vu dans la paragraphe précédent permet d'affirmer qu'un système Hamiltonien autonome avec l degrés de liberté est complètement intégrable (par quadrature) s'il possède l intégrales premières indépendantes et en involution. Nous nous proposons d'étudier dans cette section la géométrie de la variété invariante, le résultat principal est le *Théorème d'Arnol'd* que nous allons seulement énoncer, qui détermine les hypothèses sous lesquelles la variété $M_{\vec{a}}$ est un tore l -dimensionnel et donc la paramétrisation est globale, car les variables $\vec{\beta}$ résultent être des angles. Cela nous permettra ensuite (paragraphe 4) de démontrer l'existence des variables action–angle en dimension l quelconque.

THÉOREME 3.1 (Arnol'd). *Considérons un système hamiltonien autonome avec l degrés de liberté et l intégrales premières indépendantes et en involution. Si la variété de niveau $M_{\vec{a}}$ définie dans le Théorème 2.1 est compacte et connexe, alors elle est difféomorphe à un tore de dimension l .*

Observons que si $M_{\vec{a}}$ est composée de plusieurs composantes connexes le théorème s'applique à chaque composante séparément. Observons aussi qu'ils existent des systèmes qui vérifient les hypothèses du théorème de Liouville mais pas de celui d'Arnol'd, par exemple l'hamiltonien (ou tout système dont la dynamique est proche d'un point stationnaire de type selle) :

$$\mathcal{H}(p_1, p_2, q_1, q_2) = \frac{1}{2} [(p_1^2 + \omega_1^2 q_1^2) + (p_2^2 - \omega_2^2 q_2^2)] ,$$

si on pose :

$$f_1 = \frac{p_1^2 + \omega_1^2 q_1^2}{2} \quad \text{et} \quad f_2 = \frac{p_2^2 - \omega_2^2 q_2^2}{2} ,$$

alors génériquement les hypothèses du théorème de Liouville sont vérifiées, et $M_{(a_1, a_2)}$ est le produit cartésien d'une ellipse $f_1 = a_1$ dans le plan (p_1, q_1) fois une hyperbole $f_2 = a_2$ dans le plan (p_2, q_2) , donc elle n'est ni compacte ni connexe, et les hypothèses du théorème d'Arnol'd ne sont pas vérifiées.

Nous ne démontrerons pas ce résultat mais nous l'utiliserons dans le paragraphe suivant pour démontrer l'existence des coordonnées Action–Angle en dimension quelconque.

4. Systèmes hamiltoniens avec plusieurs degrés de liberté : variables action–angle

La définition des variables action–angle en dimension 1 introduite dans le paragraphe V.3 est basée sur la remarque que pour les mouvements d'oscillation chaque courbe de phase est difféomorphe à un tore. Dans le cas de systèmes hamiltoniens autonomes avec l degrés de liberté et l intégrales premières indépendantes et en involution, le théorème d'Arnol'd nous assure que si la variété de niveau est compacte et connexe alors elle est difféomorphe à un tore de dimension l . Il s'agit donc d'une situation très similaire et cela nous permettra de définir les variables action–angle dans ce cas plus général.

DÉFINITION 4.1. *Un système hamiltonien $\mathcal{H}(\vec{p}, \vec{q})$ avec l degrés de liberté est complètement canoniquement intégrable s'il existe une transformation complètement canonique de coordonnées :*

$$(VI.4.1) \quad \begin{cases} \vec{p} &= \vec{\tilde{p}}(\vec{A}, \vec{\phi}) \\ \vec{q} &= \vec{\tilde{q}}(\vec{A}, \vec{\phi}) \end{cases}$$

où la dépendance des variables $\vec{\phi}$ est $(2\pi)^l$ -périodique, des variables (\vec{p}, \vec{q}) à nouvelles variables $(\vec{A}, \vec{\phi}) \in \mathbb{R}^l \times \mathbb{T}^l$, appelées variables action–angle telles que le nouvel hamiltonien soit une fonction des seules variables d'action :

$$(VI.4.2) \quad \mathcal{H}(\vec{\tilde{p}}(\vec{A}, \vec{\phi}), \vec{\tilde{q}}(\vec{A}, \vec{\phi})) = \mathcal{K}(\vec{A}) .$$

On vérifie facilement que pour un tel système les variables d'action sont des constantes et les angles ont une évolution linéaire :

$$(VI.4.3) \quad \begin{cases} \dot{\vec{A}} &= -\nabla_{\vec{\phi}} \mathcal{K} = 0 \\ \dot{\vec{\phi}} &= \nabla_{\vec{A}} \mathcal{K} := \vec{\omega}(\vec{A}) , \end{cases}$$

dont la solution est :

$$(VI.4.4) \quad \begin{cases} \vec{A}(t) &= \vec{A}(0) \\ \vec{\phi}(t) &= \vec{\phi}(0) + \vec{\omega}(\vec{A}(0))t. \end{cases}$$

Le résultat que nous voulons démontrer est le suivant :

THÉORÈME 4.2 (variables Action–Angle dimension l). *Soit \mathcal{H} un système hamiltonien autonome avec l degrés de liberté qui possède l intégrales premières en involution et indépendantes. Supposons que pour un certain $\vec{a} \in \mathbb{R}^l$ la variété de niveau $M_{\vec{a}}$ soit compacte et connexe. Alors il existe un voisinage U de $M_{\vec{a}}$ et une transformation complètement canonique des variables $(\vec{p}, \vec{q}) \in U$ à variables action–angle $(\vec{A}, \vec{\phi}) \in V \times \mathbb{T}^l$ où V est un ouvert de \mathbb{R}^l . Le système est donc complètement canoniquement intégrable.*

Observons que grâce au théorème d’Arnol’d nous pouvons paramétriser $M_{\vec{a}}$ avec des coordonnées angles (ϕ_1, \dots, ϕ_l) , mais aussi (c’est en fait un passage de la démonstration du Théorème d’Arnol’d) d’introduire dans un voisinage de $M_{\vec{a}}$ des coordonnées *non-canoniques* $(f_1, \dots, f_l, \phi_1, \dots, \phi_l)$. Si nous fixons $\vec{f} = \vec{a}$ et tous les angles à l’exception de ϕ_j , quand il varie dans $(0, 2\pi)$ nous obtenons une courbe $\gamma_j \subset M_{\vec{a}}$. En répétant l’opération pour tous les angles nous construisons l cycles (non-homotopes dans $M_{\vec{a}}$, c’est-à-dire ils ne peuvent pas être déformés avec continuité les uns dans les autres) $\gamma_1, \dots, \gamma_l$. Cela nous permet de définir en analogie avec le cas de dimension 1 les variables d’action comme

DÉFINITION 4.3. *Les variables d’action sont définies par :*

$$(VI.4.5) \quad A_j := \frac{1}{2\pi} \int_{\gamma_j} \sum_{i=1}^l p_i dq_i \quad \text{pour } j = 1, \dots, l.$$

Apparemment cette définition est liée au choix des cycles γ_j , mais grâce à la proposition 4.4 cette ambiguïté peut être éliminée.

PROPOSITION 4.4. *Les variables action ne dépendent pas du choix des cycles γ_j (à l’intérieur de la même classe d’homotopie), si nous déformons avec continuité les cycles γ_j en des nouveaux cycles γ'_j , alors :*

$$(VI.4.6) \quad \int_{\gamma_j} \sum_{i=1}^l p_i dq_i = \int_{\gamma'_j} \sum_{i=1}^l p_i dq_i.$$

Les variables d’action dépendent seulement des fonctions \vec{f} et elles sont en involution.

DÉMONSTRATION. Le premier point est une conséquence du théorème de Stokes (voir note 8 page 43) et de la relation (VI.2.7), voir Fig. 2.

Soit donc N la variété avec bord, orientée, déterminée par les cycles γ_j et γ'_j dont la dimension est l ; le théorème de Stokes affirme que pour toute $(l-1)$ -forme θ nous avons :

$$\int_N d\theta = \int_{\partial N} \theta,$$

mais la relation (VI.2.7) nous dit que :

$$p_i dq_i|_{M_{\vec{a}}} = dW(\vec{a}, \vec{q}),$$

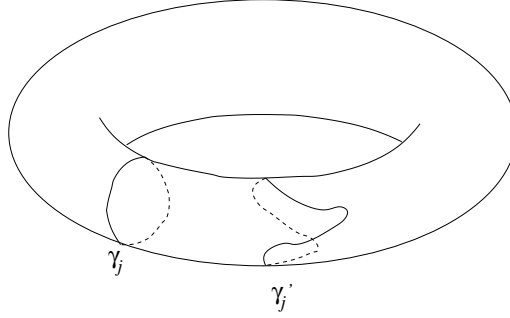


FIG. 2. Géométrie simplifiée pour démontrer l'invariance des variables d'action par rapport aux cycles γ_j .

et donc si nous prenons comme forme $\theta = p_j dq_j$, donc $d\theta = ddW = 0$ (sur $N \subset M_{\bar{a}}$) et finalement

$$0 = \int_N d\theta = \int_{\partial N} \theta = \int_{\gamma_j} p_j dq_j - \int_{\gamma'_j} p_j dq_j = 2\pi(A_j - A'_j),$$

car le bord de N est donné par les cycles γ_j et $-\gamma'_j$ (c'est-à-dire parcouru au contraire).

Par définition l'action A_j ne dépend pas de l'angle ϕ_j mais non plus des autres angles (car en changeant ces angles le cycle γ_j est déformé de façon continue et par ce qu'on vient de démontrer l'action ne change pas). Donc elles dépendent seulement des fonctions \vec{f} .

Pour montrer qu'elles sont en involution ça suffit de calculer la parenthèse de Poisson (par rapport aux variables \vec{p}, \vec{q}) de A_i et A_j ; puisque $\partial_{q_k} A_i = \partial_{f_m} A_i \partial_{q_k} f_m + \partial_{\phi_n} A_i \partial_{q_k} \phi_n = \partial_{f_m} A_i \partial_{q_k} f_m$ nous obtenons :

$$\{A_i, A_j\} = \frac{\partial A_i}{\partial f_k} \frac{\partial A_j}{\partial f_m} \{f_k, f_m\} = 0,$$

grâce à la propriété d'involution pour les fonctions \vec{f} .

Pour démontrer l'indépendance nous pouvons montrer que :

$$\det \left(\frac{\partial A_i}{\partial f_j} \right) \neq 0,$$

et en utilisant l'indépendance des fonctions \vec{f} . □

Nous pouvons finalement démontrer le Théorème 4.2.

DÉMONSTRATION. La proposition 4.4 permet d'affirmer que les variables action sont un ensemble d'intégrales premières, indépendantes et en involution pour \mathcal{H} . Pour le Théorème de Liouville la fonction

$$(VI.4.7) \quad W(\vec{A}, \vec{q}) = \int_{\vec{q}_0}^{\vec{q}} \sum_{j=1}^l p_j dq_j|_{M_{\vec{A}}},$$

est la fonction génératrice d'une transformation complètement canonique vers des nouvelles variables $(\vec{A}, \vec{\phi})$ telles que le nouvel hamiltonien est fonction des seules variables \vec{A} . Il suffit donc de montrer que les variables $\vec{\phi}$ sont des angles. Par définition

d'action A_i , l'incrément de la fonction W le long du cycle γ_i est :

$$(VI.4.8) \quad \Delta_i W = 2\pi A_i,$$

donc l'incrément de chaque variable ϕ_k sur le même cycle est :

$$(VI.4.9) \quad \Delta_i \phi_k = \Delta_i \frac{\partial W}{\partial A_k} = \frac{\partial}{\partial A_k} \Delta_i W = 2\pi \frac{\partial}{\partial A_k} A_i = 2\pi \delta_{ik}.$$

□

5. Mouvements quasi-périodiques

Dans ce dernier paragraphe nous montrons que même si intégrable, la dynamique dans les variables action-angle peut être très complexe, nous montrons aussi comme les systèmes dynamiques sont une discipline aux frontières des autres disciplines mathématiques, par exemple l'étude du mouvement pour un système complètement intégrable permet de répondre à la question suivante : quelle est le premier chiffre le plus probable dans la séquence 2^n ? Et le moins probable ? La dynamique d'un système hamiltonien peut être difficile...

L'analyse faite dans les paragraphes précédents permet d'affirmer que les mouvements des systèmes hamiltoniens complètement canoniquement intégrables sont des courbes contenues dans des variétés difféomorphes à des l -tores. Sur ces tores les orbites sont paramétrisées par :

$$(VI.5.1) \quad \begin{cases} \vec{A}(t) &= \vec{A}(0) \\ \vec{\phi}(t) &= \vec{\phi}(0) + \vec{\omega}t. \end{cases}$$

En éliminant le cas trivial $\vec{\omega} = 0$ la géométrie la plus simple est pour $l = 1$.

5.1. $l = 1$: dynamique sur \mathbb{T} . Dans ce cas la dynamique est sur un tore (cercle) de rayon $A(0)$ et avec une vitesse angulaire constante ω :

$$(VI.5.2) \quad \phi(t) = \phi(0) + \omega t \quad \text{mod } 2\pi,$$

donc toute orbite est périodique de période $T = \frac{2\pi}{\omega}$.

5.2. $l = 2$: dynamique sur \mathbb{T}^2 . Considérons maintenant le cas de dimension 2 ; un 2-tore est représenté souvent comme une chambre à air (donut en anglais) et il est le produit de deux cercles. Une façon plus intéressante de le représenter est d'utiliser un carré dont les côtés opposés sont identifiés (pensez au jeu vidéo PacMan), ainsi une orbite qui "sort du carré par le côté droit" est supposée y rentrer par le côté gauche à la même hauteur (voir Fig. 3).

Toute orbite est donc identifiée par la paramétrisation :

$$(VI.5.3) \quad \begin{cases} A_1(t) &= A_1(0) \\ A_2(t) &= A_2(0) \\ \phi_1(t) &= \phi_1(0) + \omega_1 t \quad \text{mod } 2\pi \\ \phi_2(t) &= \phi_2(0) + \omega_2 t \quad \text{mod } 2\pi, \end{cases}$$

si on élimine le temps l'orbite est décrite par une droite :

$$(VI.5.4) \quad \omega_2(\phi_1 - \phi_1(0)) - \omega_1(\phi_2 - \phi_2(0)) = 0,$$

nous pouvons supposer sans perte de généralité que la droite passe par l'origine, $\phi_1(0) = \phi_2(0) = 0$, autrement ça suffit de changer les coordonnées sur le tore pour

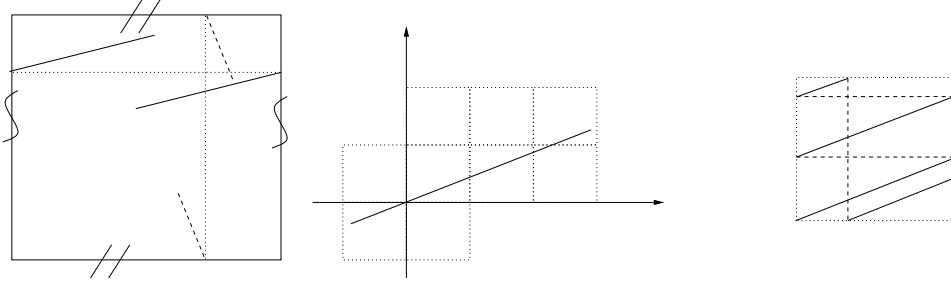


FIG. 3. Une représentation du 2-tore avec le carré dont les côtés opposés sont identifiés. Une droite sur le tore.

se placer dans cette géométrie. En réalité cette droite est contenue dans le tore et donc il faut “la couper dès qu’elle sort du carré” (voir Fig. 3).

Si $\omega_1 = 0$ (ou $\omega_2 = 0$) alors la dynamique est triviale : la droite est réduite à une droite horizontale (verticale) et les orbites sont périodiques.

Si $\omega_1\omega_2 \neq 0$ alors la situation est plus complexe ; chaque orbite intersecte l’axe verticale $\phi_1 = 0$ dans une séquence de points : $x_n = \frac{\omega_2}{\omega_1}n \bmod 2\pi$. La géométrie de cette orbite est décrite par le résultat suivant :

THÉOREME 5.1. *La séquence $x_n = \frac{\omega_2}{\omega_1}n \bmod 2\pi$ sur \mathbb{T} est périodique si et seulement si $\frac{\omega_2}{\omega_1} \in \mathbb{Q}$. Autrement si $\frac{\omega_2}{\omega_1} \in \mathbb{R} \setminus \mathbb{Q}$ alors x_n est dense dans \mathbb{T} ($B \subset C$, B est dense dans C si la fermeture de B contient C , autrement dit dans tout voisinage de n’importe quel point de C il y a un point de B .)*

DÉMONSTRATION. La séquence est périodique si et seulement si il existe un entier q tel que $q2\pi\frac{\omega_2}{\omega_1} \bmod 2\pi = 0$, c’est-à-dire il existe un deuxième entier p tel que : $q2\pi\frac{\omega_2}{\omega_1} = 2\pi p$, qui implique : $\frac{\omega_2}{\omega_1} = \frac{p}{q}$.

Si $\frac{\omega_2}{\omega_1} \in \mathbb{R} \setminus \mathbb{Q}$ alors tous les points x_n sont distincts, mais puisque le cercle est compact (borné et fermé) pour tout $\epsilon > 0$ il existe r et s entiers tels que : $|x_r - x_s| = |r2\pi\frac{\omega_2}{\omega_1} - s2\pi\frac{\omega_2}{\omega_1}| < \epsilon$. Si nous définissons $j = |r - s|$, alors la séquence finie de points $y_n = (nj2\pi\frac{\omega_2}{\omega_1} \bmod 2\pi)_{n=0}^{n=2\pi/\epsilon}$ partage le cercle en intervalles contigus de longueur plus petite que ϵ et donc tout point du cercle se trouve à une distance plus petite que ϵ d’un point de la séquence. Car ϵ est arbitraire le résultat est démontré. \square

Ce Théorème admet un corollaire important :

COROLLAIRE 5.2. *L’orbite (VI.5.4) (ou bien (VI.5.3)) dans \mathbb{T}^2 est périodique si et seulement si $\frac{\omega_2}{\omega_1} \in \mathbb{Q}$ (inclut aussi les cas $\omega_1 = 0$ ou $\omega_2 = 0$), autrement elle est dense dans \mathbb{T}^2 .*

5.3. $l \geq 3$: dynamique sur \mathbb{T}^l . En dimension $l \geq 3$ la solution de (VI.5.1) est toujours de la forme $\vec{\phi}(t) = \vec{\phi}(0) + \vec{\omega}(t - t_0)$ et on peut encore “éliminer le temps” pour l’écrire sous forme de droite dans le l -tore, $\mathbb{T}^l = \mathbb{R}^l \setminus (2\pi\mathbb{Z})^l$, c’est-à-dire le hypercube $[0, 2\pi]^l$ avec les côtés opposés identifiés.

Nous aurons toujours des orbites périodiques ou des orbites denses, mais à cause de la dimension plus grande les différents comportements “peuvent se présenter” au même temps mais dans des ensembles de dimension inférieure à l . Par exemple si

$\vec{\omega} = 2\pi(1, \sqrt{2}, \sqrt{3})$, alors l'orbite est donnée par :

$$\begin{cases} \phi_1(t) &= \phi_1(0) + 2\pi t \\ \phi_2(t) &= \phi_2(0) + 2\pi t\sqrt{2} \\ \phi_3(t) &= \phi_3(0) + 2\pi t\sqrt{3}, \end{cases}$$

et en élimant le temps nous pouvons la récrire comme (supposons $\vec{\phi}(0) = 0$) :

$$\phi_2(t) = \phi_1(t)\sqrt{2} \quad \text{et} \quad \phi_3(t) = \sqrt{3}\phi_1(t);$$

dans le “carré” qui représente le tore dans le plan (ϕ_1, ϕ_2) l'orbite a comme projection une droite de pente $\sqrt{2}$ qui donc est dense dans le carré avec les côtés identifiés, une chose similaire est valable dans le plan (ϕ_1, ϕ_3) , donc l'orbite se déroule sur un 2-tore. Une autre façon de voir cela est de considérer la séquence de points obtenus en regardant l'orbite tous les temps entiers : $x_n = \phi(n) = (2\pi n, 2\pi n\sqrt{2}, 2\pi n\sqrt{3})$. Puisque x_n appartient au tore, c'est-à-dire ses composantes sont définies modulo 2π , alors on peut réécrire $x_n = (0, 2\pi n\sqrt{2}, 2\pi n\sqrt{3})$. Donc x_n est dense dans le 2-tore : $\{0\} \times \mathbb{T}^2 \subset \mathbb{T}^3$.

Pour caractériser ce nouveau phénomène nous introduisons le *Module de résonance* (associé à $\vec{\omega}$), $\mathcal{M}_{\vec{\omega}}$:

$$(VI.5.5) \quad \mathcal{M}_{\vec{\omega}} := \{\vec{m} \in \mathbb{Z}^l \setminus \{0\} : \vec{m} \cdot \vec{\omega} = 0\}.$$

Sa dimension $d_{\vec{\omega}} = \dim \mathcal{M}_{\vec{\omega}}$ nous donne le nombre de relations de résonance indépendantes, $\vec{m} \cdot \vec{\omega} = 0$, vérifiées par $\vec{\omega}$. On peut facilement vérifier (si $\vec{\omega} \neq 0$) que $0 \leq d_{\vec{\omega}} \leq l - 1$. Par exemple dans le cas $l = 2$ nous avons deux seules possibilités : $d_{\vec{\omega}} = 0$ (aucune relation de résonance) c'est-à-dire pour tout $\vec{m} \in \mathbb{Z}^2 \setminus \{0\}$ nous avons $\vec{\omega} \cdot \vec{m} \neq 0$, donc :

$$0 \neq m_1\omega_1 + m_2\omega_2 \Rightarrow \frac{\omega_1}{\omega_2} \neq \frac{m_2}{m_1} \Rightarrow \frac{\omega_1}{\omega_2} \notin \mathbb{Q}.$$

Ou bien $l = 1$, (cas complètement résonant), il existe $\vec{m} \in \mathbb{Z}^2 \setminus \{0\}$ t.q. $\vec{\omega} \cdot \vec{m} = 0$, donc :

$$0 = m_1\omega_1 + m_2\omega_2 \Rightarrow \frac{\omega_1}{\omega_2} = \frac{m_2}{m_1} \Rightarrow \frac{\omega_1}{\omega_2} \in \mathbb{Q}.$$

Le cas $l = 2$ ne peut pas se produire car on aurait :

$$\exists \vec{m}, \vec{n} \in \mathbb{Z}^2 \setminus \{0\} : \vec{m} \neq \vec{n}, \vec{\omega} \cdot \vec{m} = \vec{\omega} \cdot \vec{n} \Rightarrow \vec{\omega} \cdot (\vec{m} - \vec{n}) = 0,$$

et puisque par hypothèse $\vec{m} \neq \vec{n}$ la seule possibilité serait d'avoir $\vec{\omega} = 0$.

Le Corollaire 5.2 peut se relire comme : si $d_{\vec{\omega}} = 0$ alors l'orbite est dense dans un 2-tore et si $d_{\vec{\omega}} = 1$ alors l'orbite est périodique dans un 2-tore ; ce résultat peut être généralisé en dimension quelconque :

THÉOREME 5.3. Soit $\mathcal{M}_{\vec{\omega}}$ le module de résonance associé à la fréquence $\vec{\omega}$, alors :

- (1) l'orbite est périodique si et seulement si $d_{\vec{\omega}} = l - 1$ (résonance complète) ;
- (2) si $d_{\vec{\omega}} = 0$ (absence de résonances) l'orbite est dense dans un l -tore ;
- (3) si $0 < d_{\vec{\omega}} < l - 1$ alors l'orbite est dense dans un tore de dimension $l - d_{\vec{\omega}}$ immergé dans \mathbb{T}^l .

Les orbites correspondantes aux cas (2) et (3) sont appelées *quasi-périodiques*. De façon générale nous pouvons définir :

DÉFINITION 5.4. Une fonction continue $g : \mathbb{R} \rightarrow \mathbb{R}$ est dite quasi-périodique s'il existe un vecteur $\vec{\omega} = (\omega_1, \dots, \omega_l) \in \mathbb{R}^l$ et une fonction continue $f : \mathbb{T}^l \rightarrow \mathbb{R}$ tels que :

$$(VI.5.6) \quad g(t) = f(\omega_1 t, \dots, \omega_l t).$$

REMARQUE 5.5. Le nom quasi-périodique est dû au fait que la fonction se comporte comme si elle était périodique mais avec des périodes de plus en plus longues. Par exemple si $l = 2$ et si $\vec{\omega} = 2\pi(1, \sqrt{2})$, alors d'après le cours d'analyse on sait qu'il existe une séquence de rationnels (p_k, q_k) qui converge vers $\sqrt{2}$, de plus cette séquence peut être choisie de façon que $|\sqrt{2} - \frac{p_k}{q_k}| \leq \frac{1}{q_k q_{k+1}}$. Si nous avons une fonction quasi-périodique telle que $g(t) = f(t, t\sqrt{2})$, pour une certaine fonction f définie sur le 2-tore, alors c'est facile de vérifier que pour tout $t \in [0, q_k]$:

$$\begin{aligned} |g(t + q_k) - g(t)| &= |f(2\pi t + 2\pi q_k, 2\pi t\sqrt{2} + 2\pi\sqrt{2}q_k) - f(2\pi t, 2\pi t\sqrt{2})| \\ &= |f(2\pi t, 2\pi t\sqrt{2} + 2\pi\sqrt{2}q_k - 2\pi p_k) - f(2\pi t, 2\pi t\sqrt{2})| \\ &\leq C 2\pi |\sqrt{2}q_k - p_k| \leq 2\pi C \frac{1}{q_{k+1}}, \end{aligned}$$

où nous avons utilisé au deuxième passage la périodicité de f sur le 2-tore et dans le dernier passage la propriété de la séquence p_k/q_k . Donc si k est grand, puisque $q_k \rightarrow \infty$, $|g(t + q_k) - g(t)|$ est petit. Il serait égal à zéro si g était périodique.

La moyenne temporelle d'une fonction quasi-périodique est définie comme :

$$(VI.5.7) \quad \langle g \rangle_T := \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T g(t) dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(\omega_1 t, \dots, \omega_l t) dt,$$

si cette limite existe.

A côté de la moyenne temporelle nous pouvons considérer la *moyenne en phase* (spatiale) de la fonction f :

$$(VI.5.8) \quad \langle f \rangle := \frac{1}{(2\pi)^l} \int_{\mathbb{T}^l} f(\vec{\phi}) d\vec{\phi}.$$

Considérons maintenant "l'orbite" d'une fonction quasi-périodique, $\cup_{t \in \mathbb{R}} g(t)$; puisque l'orbite peut être dense dans le tore, nous pouvons comparer les deux moyennes et pouvoir déterminer si l'orbite est dense ou pas. En gros si l'orbite est dense, le point $g(t)$ décrit presque tout le tore et donc sa moyenne temporelle sera égale à sa moyenne spatiale ; plus précisément :

THÉORÈME 5.6. Soit $f : \mathbb{T}^l \rightarrow \mathbb{R}$ une fonction continue, et considérons la fonction quasi-périodique obtenue en composant le flot (VI.5.1) avec f , $g(t) = f \circ \vec{\phi}(t)$. Si la fréquence $\vec{\omega}$ n'est pas résonante, i.e. $\dim \mathcal{M}_{\vec{\omega}} = 0$, alors la moyenne temporelle $\langle g \rangle_T$ existe, est constante et coïncide avec la moyenne spatiale.

DÉMONSTRATION. Montrons d'abord le théorème dans le cas où f est un polynôme trigonométrique, somme finie de fonctions sin et cos, pour simplifier les notations nous l'écrirons avec les exponentielles complexes :

$$(VI.5.9) \quad f(\vec{\phi}) = \sum_{\vec{m} \in \mathcal{F}} \hat{f}_{\vec{m}} e^{i\vec{m} \cdot \vec{\phi}},$$

où $\mathcal{F} \subset \mathbb{Z}^l$ est un ensemble fini d'indices de sommation. La fonction quasi-périodique g aura une décomposition similaire :

$$g(t) = \sum_{\vec{m} \in \mathcal{F}} \hat{f}_{\vec{m}} e^{i\vec{m} \cdot (\vec{\phi}_0 + t\vec{\omega})}.$$

Prenons maintenant un élément $\vec{m} \in \mathcal{F}$ et considérons sa contribution à la moyenne temporelle de g , puisque la moyenne est une intégrale elle résulte linéaire, donc la moyenne de g sera la somme des contributions de chaque indice \vec{m} . Si $\vec{m} = 0$ alors $g_0 = \hat{f}_0$ est une constante et donc

$$\langle g_0 \rangle_{T=\infty} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \hat{f}_0 dt = \hat{f}_0 = \frac{1}{(2\pi)^l} \int_{\mathbb{T}^l} \hat{f}_0 d\vec{\phi}.$$

Si $\vec{m} \neq 0$ alors la contribution à la moyenne spatiale vaut zéro :

$$\begin{aligned} \langle \hat{f}_{\vec{m}} e^{i\vec{m} \cdot \vec{\phi}} \rangle &= \frac{1}{(2\pi)^l} \int_{\mathbb{T}^l} \hat{f}_{\vec{m}} e^{i\vec{m} \cdot \vec{\phi}} d\vec{\phi} = \frac{\hat{f}_{\vec{m}}}{(2\pi)^l} \int_0^{2\pi} e^{im_1 \phi_1} d\phi_1 \dots \int_0^{2\pi} e^{im_l \phi_l} d\phi_l \\ &= \frac{\hat{f}_{\vec{m}}}{(2\pi)^l} \frac{e^{im_1 \phi_1}}{m_1} \Big|_0^{2\pi} \dots \frac{e^{im_l \phi_l}}{m_l} \Big|_0^{2\pi} = 0. \end{aligned}$$

Par ailleurs la contribution à la moyenne temporelle vaut :

$$\begin{aligned} \langle g_{\vec{m}} \rangle &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \hat{f}_{\vec{m}} e^{i\vec{m} \cdot (\vec{\phi}_0 + \vec{\omega}t)} dt = \lim_{T \rightarrow \infty} \frac{\hat{f}_{\vec{m}} e^{i\vec{m} \cdot \vec{\phi}_0}}{T} \int_0^T e^{i\vec{m} \cdot \vec{\omega}t} dt \\ &= \lim_{T \rightarrow \infty} \frac{\hat{f}_{\vec{m}} e^{i\vec{m} \cdot \vec{\phi}_0}}{T} \frac{e^{i\vec{m} \cdot \vec{\omega}T} - 1}{i\vec{m} \cdot \vec{\omega}} \Big|_0^T = \lim_{T \rightarrow \infty} \frac{\hat{f}_{\vec{m}} e^{i\vec{m} \cdot \vec{\phi}_0}}{T} \frac{e^{i\vec{m} \cdot \vec{\omega}T} - 1}{i\vec{m} \cdot \vec{\omega}} = 0, \end{aligned}$$

et l'intégration est bien définie car $\vec{m} \cdot \vec{\omega} \neq 0$. Donc les contributions aux deux moyennes sont égales ; par linéarité des intégrales et puisque f est composé d'un nombre fini de termes, nous pouvons démontrer que la moyenne spatiale et la moyenne temporelle sont égales.

Prenons maintenant une fonction f continue quelconque sur \mathbb{T}^l , alors par le Théorème de Weierstrass nous pouvons donner une approximation à ϵ près pour tout $\epsilon > 0$, avec un polynôme trigonométrique $P_\epsilon(\vec{\phi})$:

$$\max_{\vec{\phi} \in \mathbb{T}^l} |f(\vec{\phi}) - P_\epsilon(\vec{\phi})| < \frac{\epsilon}{2}.$$

Soit $P_\pm := P_\epsilon \pm \frac{\epsilon}{2}$, alors ¹ $P_- \leq f \leq P_+$ et :

$$\frac{1}{(2\pi)^l} \int_{\mathbb{T}^l} (P_+(\vec{\phi}) - P_-(\vec{\phi})) d\vec{\phi} = \epsilon.$$

Donc pour tout $\epsilon > 0$ il existe deux polynômes trigonométriques P_+ et P_- tels que ² :

$$\langle f \rangle - \langle P_- \rangle \leq \epsilon \quad \text{et} \quad \langle P_+ \rangle - \langle f \rangle \leq \epsilon,$$

¹En effet pour tout $\vec{\phi} \in \mathbb{T}^l$ nous avons : $f(\vec{\phi}) - P_-(\vec{\phi}) = f(\vec{\phi}) - P_\epsilon(\vec{\phi}) + \frac{\epsilon}{2}$; or par définition de P_ϵ nous avons $|f(\vec{\phi}) - P_\epsilon(\vec{\phi})| \leq \epsilon/2$ pour tout $\vec{\phi} \in \mathbb{T}^l$, donc en particulier $f(\vec{\phi}) - P_\epsilon(\vec{\phi}) \geq -\epsilon/2$, et finalement $f(\vec{\phi}) - P_-(\vec{\phi}) \geq 0$.

²Pour monter cela il suffit d'utiliser la définition de moyenne spatiale :

$$\langle f \rangle - \langle P_- \rangle = \frac{1}{(2\pi)^l} \int_{\mathbb{T}^l} (f(\vec{\phi}) - P_-(\vec{\phi})) d\vec{\phi} \leq \frac{1}{(2\pi)^l} \int_{\mathbb{T}^l} (f(\vec{\phi}) - P_\epsilon(\vec{\phi}) + \frac{\epsilon}{2}) d\vec{\phi} \leq \frac{\epsilon}{2} + \frac{\epsilon}{2}.$$

et pour tout T :

$$\frac{1}{T} \int_0^T P_- \circ \vec{\phi}(t) dt \leq \frac{1}{T} \int_0^T f \circ \vec{\phi}(t) dt \leq \frac{1}{T} \int_0^T P_+ \circ \vec{\phi}(t) dt.$$

Nous avons déjà démontré que la moyenne temporelle d'un polynôme trigonométrique est zéro, donc pour tout $\epsilon > 0$ il existe $T(\epsilon)$ tel que :

$$| \langle P_{\pm} \rangle - \frac{1}{T(\epsilon)} \int_0^{T(\epsilon)} P_{\pm} \circ \vec{\phi}(t) dt | \leq \epsilon,$$

mais alors en mettant tout ensemble nous avons démontré que pour tout $\epsilon > 0$, il existe $T(\epsilon)$ tel que pour tout $T > T(\epsilon)$ nous avons :

$$| \langle f \rangle - \frac{1}{T} \int_0^T f \circ \vec{\phi}(t) dt | \leq 2\epsilon,$$

et puisque ϵ est arbitraire cela conclut la démonstration. \square

REMARQUE 5.7. Avec ce résultat nous pouvons facilement montrer la deuxième propriété du Théorème 5.3. Prouvons ce résultat par contradiction. Supposons que $\dim \mathcal{M}_{\vec{\omega}} = 0$ et supposons que l'orbite ne soit pas dense dans \mathbb{T}^l ; il existe alors un point $\vec{\phi} \in \mathbb{T}^l$ et un voisinage ouvert, U , tel que l'orbite n'intersecte jamais cet ouvert. Prenons maintenant une fonction continue quelconque $f : \mathbb{T}^l \rightarrow \mathbb{R}$ avec les propriétés suivantes :

- (1) $\langle f \rangle = 1$;
- (2) $f(\vec{\phi}) = 0$ pour tout $\vec{\phi} \notin U$.

Mais alors la moyenne temporelle serait égale à zéro, par contre sa moyenne spatiale serait égale à 1, ce qui contredit le Théorème 5.6. Donc l'orbite doit être dense et cet ouvert U n'existe pas.

Nous terminons ce Chapitre avec l'application suivante de la Théorie des fonctions quasi-périodiques.

EXEMPLE 5.8 (Arnol'd). Considérons la séquence obtenue en prenant les premiers chiffres de la séquence 2^n :

$$a_n := \text{premier chiffre de } 2^n,$$

donc

$$a_0 = 1, a_1 = 2, a_3 = 8, a_4 = 1 \dots$$

Nous voulons répondre à la question suivante : est-il plus facile de trouver $a_n = 1$ ou $a_n = 5$? Pour cela nous allons calculer la probabilité que a_n soit un entier, k , (entre 1 et 9) donné. Or $a_n = k$ si et seulement si :

$$\log_{10} k \leq \{n \log_{10} 2\} < \log_{10}(k+1),$$

où $\{x\}$ est la partie fractionnaire de x . En effet si nous écrivons

$$2^n = s, \dots 10^m,$$

avec $1 \leq s < 10$, alors $a_n = s$, et

$$\{\log_{10} 2^n\} = \{m + \log_{10}(s, \dots)\} = \log_{10}(s, \dots),$$

mais alors si ($s < 10$ et donc $\log_{10}(s, \dots) < 1$) :

$$\log_{10} k \leq \{\log_{10} 2^n\} = \log_{10}(s, \dots) < \log_{10}(k+1),$$

nous avons :

$$k \leq s, \dots < k+1,$$

c'est-à-dire k est l'entier que nous cherchons.

Le nombre $\log_{10} 2$ est irrationnel et donc par le Théorème 5.1 la séquence $\{n \log_{10} 2\}$ est dense dans l'intervalle $[0, 1]$.

Or la fréquence (probabilité) avec laquelle l'entier k paraît dans a_n est donnée par :

$$(VI.5.10) \quad p_k := \lim_{N \rightarrow \infty} \frac{\#\{\{n \log_{10} 2\} \in [\log_{10} k, \log_{10}(k+1)) : 0 \leq n \leq N-1\}}{N},$$

où $\#A$ est la cardinalité de l'ensemble A (nombre d'éléments contenus).

Si nous considérons la fonction $\chi_k : [0, 1] \rightarrow \mathbb{R}$:

$$\chi_k(x) = \begin{cases} 1 & \text{si } x \in [\log_{10} k, \log_{10}(k+1)) \\ 0 & \text{autrement,} \end{cases}$$

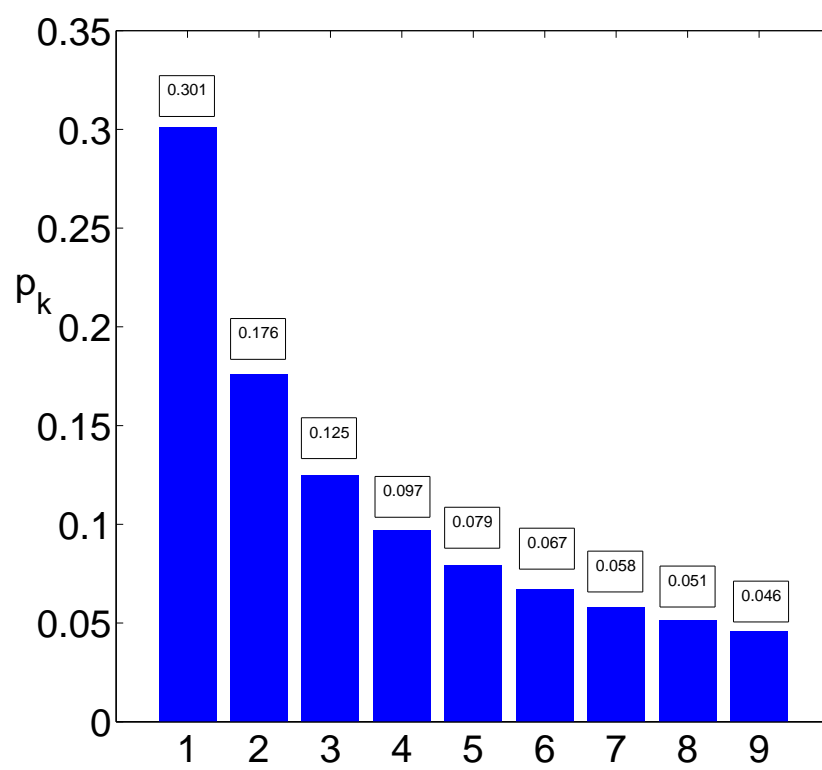
alors la (VI.5.10) est la moyenne temporelle (ici nous considérons la variable temporelle discrète et non continue, mais cela ne change pas les résultats précédents) de la fonction χ_k calculée le long de la séquence $\{n \log_{10} 2\}$:

$$(VI.5.11) \quad \langle \chi_k \rangle_T := \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{j=0}^{T-1} \chi_k(\{n \log_{10} 2\}) = p_k.$$

Or nous avons démontré le Théorème 5.6 sous l'hypothèse de continuité pour la fonction f , mais en réalité il est valable pour une fonction discontinue (l'approximation de Weierstrass sera ponctuelle et non uniforme), donc la moyenne temporelle de χ_k est égale à sa moyenne spatiale :

$$(VI.5.12) \quad p_k = \langle \chi_k \rangle := \int_0^1 \chi_k(x) dx = \log_{10}(k+1) - \log_{10} k.$$

Comme on montre dans la Fig. 4 le chiffre le plus probable est 1 avec environ 30 % de probabilité, et les autres suivent en ordre décroissant.

FIG. 4. Probabilités p_k en fonction de l'entier k .

CHAPITRE VII

Bibliographie.

Les sujets traités dans ce cours peuvent être retrouvés dans beaucoup de livres *Classiques de mécanique hamiltonienne* ou *mécanique analytique*, nous proposons une liste (forcément incomplète) qui pourra aider l'étudiant à compléter sa préparation. La plupart de ces livres sont présents dans la bibliothèque de l'unité de systèmes dynamiques.

- R. Abraham et J.E. Marsden : *Foundations of mechanics*. (Chapitres 2,3 et 6)
- V. Arnol'd : *Méthodes mathématiques de la mécanique classique*
- A. Fasano et S. Marmi : *Analytical Mechanics*. (Tout le livre)
- J.V. José et E.J. Saletan : *Classical Dynamics*. (Chapitres 5 et 6)
- L. Meirovitch : *Methods of analytical dynamics*. (Chapitres 2 et 9)
- J. Roels : *Mécanique lagrangienne et hamiltonienne*. (Tout le livre)